A Compendium of English Orthography

Items in blue are hyperlinks to the appropriate sections of the compendium and to files elsewhere on the site. One easy way to return to the location of the original hyperlink is to use the page thumbnails, which can be revealed by pressing F4.

A, an
Ablaut, umlaut
-able), -ible)
Accede, exceed, proceed, succeed
Access, excess
Accuse, excuse
Adapt, adopt
Addition, edition
Adjectives, regular and nonregular
Adoptions and adoptions
Adverbs
Advise, advice
Affect, effect
Affixes
Affluent, effluent
Affricate sounds
Alfred the Great
All and its compounds
Alley, ally
Alliteration
Alveolar sounds
American Sign Language (ASL)
Analogy
-ance), -ence); -ant), -ent)
Angel, angle
Angles, Saxons, Jutes, and Frisians
Anglo-Saxon, or Old English
Apostrophe
Artificial
Assimilation
Assure, ensure, insure
Back formation
Bases
Believe, belief
Beneficial, beneficiary
Beowulf
Bilabial sounds
Blends
Braille
British and American spelling
Capital and lowercase letters
Capital, capitol
Cardinal and ordinal numbers
Case
Caxton, William
Changes in Some Indo-European Sounds
Changing <y> to <i> and <i> to <y>
Chaucer, Geoffrey
Clauses and sentences
Closed syllables
Code and performance
Comparative and superlative
Compliment, complement
Compound words
Concatenation
Concatenations
Concrete vs. abstract, and specific vs. general
Confident, confidant
Conjunctions
Conscious, conscience
Consonant and vowel letters
Consonant letters and the sounds they spell
Consonant sounds
Content and meaning
Contraction
Conversions
Council, counsel, consul, consult
CVC# strings
Decent, descent, dissent
Deficit
Dental sounds
Derivation, derived forms
Desert, dessert, deserts
Diacritic marks
Dialects
Dictionaries
Digraphs and trigraphs
Digraphs <ch, gh, ph, rh, sh, th, wh>
Diphthongs
Doublets and doublet equivalents
Doublets misspelled
Early Modern English
*Efficient, efficiency*
Elements
*Eminent, imminent*
*Envelope, envelop*
-er)*01, -or)*2
*Essay, assay*
Etymology
Evolution
Explication
*Excel, excellent*
*Except, accept*
*Finally, finely*
Folk etymology
Fossils
French influence
French Lemon (Stress Frontshift) Rule
Frequentatives
Fricative sounds
Futhorc
Genitive
*Germanic languages*
Glides
Glottal sounds
Graphs and graphemes
Great Vowel Shift
Holdouts
Homonyms, homophones, homographs
Hyphen
<\#> before <E>
*Illusion, allusion*
Imitative
*Indo-European languages*
*Indo-European roots*
Inflection, inflected forms
*Innocence, innocents*
Insulators
Intensifiers
Interdental sounds
International Phonetic Alphabet (IPA)
*Its, it’s*
-ize), -ise)
Jespersen, Otto
Johnson, Samuel
Labiodental sounds
Language families
Latin alphabet
Latinization
Lie, lay
Long vowels
Lose, loose
Medal, metal, mettle, meddle
Metaphor
Metonymy
Middle English
Minims
Missionaries
Misspellings
Mnemonic
Modern English
Morphemes and Morphophonemics
Mulcaster, Richard
Nasal sounds
Nonce words
Norman Conquest
Norman scribes
Norse and Anglo-Norse
Noun plurals
Nouns
Number Words
Office, officer, official
Old English, or Anglo-Saxon
Onomatopoeia
Open and closed syllables
Oxford English Dictionary (OED)
Palatalization
Palatal sounds and their spellings
Participles
Particles
Partridge, Eric
Pastor, pasture
(Per-, (pre-
Periphrastic
Personal, personnel
Phonemics vs. phonetics
Phonetic spelling
Plurals, regular and nonregular
Syllabic consonants
Syllables
Synonyms, homonyms, antonyms, pseudonyms, heteronyms, acronyms
Third Syllable (or Third Vowel) Rule
Through, thorough
Triplets
Twinning
Use, used
Variant spellings
VCCle vs. VCe strings
VCV vs. VCCV strings
VCV vs. VCCV contrasts
Velar sounds
Verbs: regular and irregular
Vikings
Voiced vs voiceless (unvoiced) sounds
Vowel lengthening and shortening
Vowel letters
Vowel sounds
V.V strings
Wear, ware, where; were, we’re
Webster, Noah
Wessex and the West Saxon dialect

Further sources: The sources of further information are limited for the most part to three: (i) David Crystal's An Encyclopedic Dictionary of Language and Languages (Blackwell, 1992) (hereafter EDLL), (ii) The Oxford Companion to the English Language, ed. Tom McArthur (Oxford, 1992) (OCEL), and (iii) my American English Spelling (Johns Hopkins, 1988) (AES). EDLL and OCEL are arranged alphabetically with headings the same as or close to those here. All three often contain references to sources of further information.

Symbols Used: Letters and spellings are enclosed in arrowhead brackets: <a>, <cat>. Sounds and pronunciations are enclosed in square brackets: [ă], [kăt]. Definitions are enclosed in double quotes. Prefixes are marked with a left parenthesis and trailing hyphen: (pre-, (non-; suffixes with a leading hyphen and a right parenthesis: -ing)1, -ed)1. The unstressed, reduced sound schwa is represented with [ə]. The analyses, or explications, of written words are underlined, with element boundaries marked with plus signs: catfish cat1+fish1 and deletions with an overstruck diagonal: academe+ic1. In phonetic respellings primary word stress is indicated with
A following high vertical, secondary with a low vertical, and syllable boundaries that have no stress marks are indicated with mid dots: [flæˈmɪŋ], [bɪɡˌhɛdˈɪd], [əˈbʊv]. The index numbers on some prefixes, bases, suffixes, and particles – for instance, (in2-, -ed)1, ag5 – identify members of sets of homographs. To determine which element or particle is being referred to, consult the Prefixes, Bases, Suffixes, or Particles datatable in the Lexis database on this site. Misspellings are marked with a leading asterisk – *mispellings.

A, an; the. A and an are indefinite articles, normally used to refer to something being mentioned for the first time, while the definite article the is normally used to refer to things that were mentioned earlier or that can be assumed. A is used before nouns that start with a consonant sound, and an is used before nouns that start with a vowel sound – a berry but an orange. The important thing is the first sound of the noun, not necessarily the first letter – thus, it’s an hour (with the initial silent <h>), not *a hour, and a unit (with the initial [y] sound), not *an unit. See <h> and <u>.

Ablaut, umlaut. In German laut means “sound” and is related to our word loud; the German ab means “off” and um means “alteration.” Ablaut means a change in vowel sound, often to indicate a change in grammatical function. It occurs in several Indo-European languages, including English, as in man, men; goose, geese; mouse, mice; sing, sang, sung. See Strong verbs and Plural nouns. Umlaut has two senses: (i) a modification in a vowel sound, and (ii) a mark, sometimes called dieresis or diaeresis, consisting of two dots placed above a vowel to show that modification, as in German singular mann “man” vs. its plural männer “men”, reflected in English man vs. men. An umlaut, or dieresis, is also sometimes used to indicate that a vowel is not silent, as in the surname Brontë.

-able), -ible). Since these two synonymous suffixes are homophones, it can be hard to know which one to use. Although there are no reliable rules for knowing when to pick one over the other, these three observations can help: (i) The suffix -able is about six times more common than -ible, so when in doubt, choose -able). (ii) Verbs that end with -ate)1 and/or form nouns with -ation) take -able) – imitate, imitable; civilize, civilization, civilizable. And (iii) since -able) and -ible) come late in their words, the correct spelling is easy to find in a dictionary. Two other points, though neither raises any problems for spellers: A small, and shrinking, number of words can take either -able) or -ible) – as in ascendable, ascendible. And a very small number of words with -able) have forms with and without final
**<e> deletion**: likeable, likable; liveable, livable; moveable, movable; sizeable, sizable. This confusion is probably due to <able> sometimes being seen not as the suffix -able but as the etymologically unrelated word able. With able the word formed would be a compound, and compounds don’t require final <e> deletion.

**Accent.** Accent has two senses important to us: The first is a characteristic or typical pronunciation, as in “a Brooklyn accent” or “a Spanish accent.” The second is stress.

**Adapt, adopt.** The prefix (ad-) carries the sense “to, toward.” The base apt1 means “to fit” and the base opt2 means “to choose.” So adapt (ad+apt1) means “to fit something to something else,” and adopt (ad+opt2) means “to choose something.”

**Addition, edition.** The prefix (ad-) carries the sense “to, toward.” The prefix (e-, a contraction of (ex-, carries the sense “away, out.” The base dit1 means “give, put, bring.” The suffix -ion)1 forms nouns. Addition (ad+dit1+ion)1 thus means something like “a giving to,” and edition (e+dit1+ion)1 means “a putting or bringing out.”

**Adjectives, regular and nonregular.** Adjectives are usually defined as words that modify a noun. Another kind of definition is “Any word that does not end in -’s) but will fit into this blank: The _______ thing seemed okay.”

**Regular** adjectives can take the comparative and superlative inflectional suffixes -er, -est – as in dark, darker, darkest. But in many, or all, cases the comparative and superlative can also be shown periphrastically with more and most – more dark, most dark. There are three kinds of nonregular adjectives:

- those that show comparative and superlative only periphrastically, as in admirable, more admirable, most admirable but not *admirabler, *admirablest;
- those that have comparative and superlative forms with bases different from the positive form, as good, better, best; bad, worse, worst; far, further, furthest; and
- those that only rarely or never have comparative or superlative forms – for instance, ordinals like eighteenth; possessive adjectives like her, his, your, my, our; and certain absolutes like every, subsequent, prior.

**Adoptions and adoptions.** With words, I prefer the notion of adopting rather than borrowing, since when you borrow something, you intend to return it, but when you adopt something, you intend to keep it, as we do when we adopt words from other languages. Many words are adopted
without any change in spelling – *poi, kangaroo, scaloppini, soufflé, and the like. But usually we make changes, often minor, sometimes major, in spelling. These I call adaptions. A small sample of the tens of thousands in English, with the language from which each was adapted:

- block < French *bloc
- buckaroo < Spanish *vaquero
- budget < Old French *bougette
- cashew < Portuguese *acajú
- cream < Old French *cresme
- hoosegow < Spanish *juzgado
- jerky < Spanish *charqui
- jolly < Old French *jolif
- ransack < Old Norse *rannsaka

See *Evolution* and AES, pp. 17-20.

**Adverbs.** Adverbs come in different shapes and have several different functions. They can modify:

- **verbs:** He spoke **loudly**.
- **adjectives:** He is **extremely** loud.
- other adverbs: He spoke **very** loudly.
- an entire clause or sentence: **Truthfully**, he’s a loudmouth.

When they modify adjectives or adverbs, adverbs are usually **intensifiers**, like **extremely**. Most adverbs end with the **derivational suffix** -ly1 which turns adjectives into adverbs, as in **loudly**, **extremely**, and **truthfully**. Like adjectives, most one-syllable adverbs form their comparative and superlative with -er02 and -est1: slow, slower, slowest. But adverbs two or more syllables long form their comparative and superlative **periphrastically** with more and most: loudly, more loudly, most loudly.

**Advise, advice.** The prefix (ad- means “to, toward.” The bases *vice*3, which forms nouns, and *vise*2, which forms verbs, originally meant “see, look,” but in Latin they developed a meaning like “my view or opinion is.” Remember that <s> often spells the sound [z] as it does in *advise*, but <c> always spells [s] before <e> as it does in *advice*.

**Affect, effect.** These two **homophones** (and near **homographs**) are found on most lists of **spelling demons**. Usually **affect** is a **verb** – as in “Caffeine affects a lot of people.” And usually **effect** is a **noun** – as in “Caffeine doesn’t have much of an effect on me.” To affect is to cause while an effect is the result. There may be some help in remembering that there’s an <a> and an <e> in both **affect** and **cause**. However, an affect “feeling, emotion” is also a noun – as in “The defendant revealed no affect to the
jury’s verdict”; and *effect* can be a verb – as in “The new owner will effect a total change in the company.” Mercifully, these final two uses are quite rare: *affect* as a noun occurs mainly in the psychological *register*, and *effect* as a verb tends to be rather formal.

**All and Its Compounds.** The bound base *all05* is a *contraction* of *all1* “all, entire” – as in *almighty, alone, also, although, almost, wherewithal*. But some words with *all05* have partners with *all1*, which can lead to spelling problems. For instance, we have *always* “invariably, forever” – as in “She’s always late” and *all ways* “in each and every way” – as in “I’ve tested it in all ways.” Other examples: *already* “by this time, so soon” vs. *all ready* “each and every one is prepared”; the adverb *altogether* “entirely” vs. the adjective phrase *all together* “in a single group.” And this more sneaky pair: the adjective phrase *all right* vs. the common but still nonstandard *alright*. See [Excel, excellent](#).

**Affixes.** Affixes are bound elements that can be added to *stems* to change their function or sense. In English the two kinds of affixes are *prefixes* and *suffixes*.

**Affricate sounds.** Affricates are *consonant sounds* that begin with a *stop* and end with a *fricative* – for example, the affricates *[j] = [dzh], [ch] = [tsh]*, as in *judge* and *church*.

**Alfred the Great.** From 871-899 Alfred was the Christian king of Wessex, one of the seven kingdoms of early England, located south and west of the Thames River. He lived from 849-899. While king he reorganized military strategy and directed the defeat of the early Viking raiders. The dialect of Wessex was West Saxon, and Alfred also established a school in his court that taught both Latin and the West Saxon dialect of *Old English*. To further education in his kingdom, he had several important works translated from Latin into West Saxon, he himself translating some. His emphasis on translation into English helped make West Saxon very important in the development of and our understanding of early English.

**Alley, ally.** These two words are near *homographs* though different in pronunciation – *[ālˈ e] and *[ālˈ ē]* respectively. *Alley* explicates to *all4+ey)2*. The base *all4* means “go, walk” and the suffix *-ey)2* forms nouns. So an alley is where one walks or goes to some place. *Ally* explicates to *(ad+I+ly2*. The prefix *(ad-* means “to, toward,” and the base *ly2* means “bind, join,” so an ally is one to whom you are bound or joined.

**Alliteration.** Alliteration is the repetition of sounds – usually *consonant* sounds – at the beginning of words and in *stressed syllables*. Alliteration is
very common in poetry (as in Wallace Stevens’ “\textit{W}inding across \textbf{w}ide \textit{w}ater, without sound./ The day is like \textbf{w}ide \textit{w}ater, without sound.” It was the basis of \textit{Old English} poetry, as in \textit{Beowulf}. It is also common in popular phrases like “dead as a doornail.”

\textbf{Alveolar sounds.} Alveolars, or dentals, are \textit{consonant sounds} articulated with the front of the tongue near or touching the alveolar ridge behind the upper teeth – [d, t, z, s, n, l, r].

\textbf{American Sign Language (ASL).} ASL is a visual language used by probably 250,000 to 500,000 people – people who are hard of hearing or totally deaf, or live or work with people who are. ASL was developed at the American School for the Deaf in the early 19\textsuperscript{th} century by Thomas Gaulladet. While in France at the national institute in Paris, Gaulladet had studied French methods for instructing the deaf. He returned to the United States with Laurent Clerc from that institute, and together they developed ASL, basing it on the French model, together with signing practices from local American families and communities.

One estimate is that ASL involves an astonishing 360,000 physical signs, involving mostly the hands but also facial expressions and body movements. These signs can be described using five parameters: (i) the shape of the hands, (ii) their movement, (iii) the orientation of the palms, (iv) the hands’ location, plus (v) non-manual markers involving the eyebrows, the cheeks, the nose, the head, the torso, the mouth and lips, and the eyes.

ASL is a natural language with its own grammar and syntax. Its signs function like the \textit{phonemes} and \textit{morphemes} of a spoken language. It has different \textit{dialects}, and even within a dialect signers have their own idiosyncratic styles of presentation. As ASL has gained recognition as a natural language, more and more colleges have come to accept it for foreign language credit.

ASL is not a universal sign language. For instance, America, Britain, and Australia, three English-speaking nations, have three different sign languages that are not mutually intelligible – just as in \textit{Middle English} the spelling of speakers of northern British dialects was often quite unintelligible to readers from the south, and the way that today the varieties of spoken English from different parts of the world can be difficult and even unintelligible to native speakers of American English.

One of the features of ASL is \textit{fingerspelling}, used to spell out words for which there is no known sign. Fingerspelling involves manual signs for the
Several websites provide videos and explanations of many of ASL’s signs – for instance, American Sign Language University at http://www.lifeprint.com/.

**Analogy.** In the most general sense, analogy is the perception of similarity in the midst of difference. Analogy is perhaps the most important way in which the mind searches for unity in the din and roar of experience, linguistic and otherwise. It underlies metaphor. In spelling and vocabulary it leads to users changing the sound or spelling or both of a word to make it more similar to a more regular or more familiar form—as when French **carriole** “a small covered carriage” is changed by folk etymology to **carryall** or when Middle English **coutes, coude** is respelled to **could**, making it more similar to the parallel forms **should** and **would** (which in Middle English were **scholde** and **wolde**). (EDLL, OCEL, AES, pp. 10-12)
The suffixes -ance and -ence can create problems because they are homophones pronounced [əns] and both form nouns, usually abstract. There are no simple rules for knowing when to use one and when the other, but these two hints may help some:

• If you can add [ensh] (spelled <ential>) to the stem and get a recognizable word, the [əns] is -ence. For instance, if you can't decide between <confidence> and <confidance>, and you replace the [əns] with [ensh], the result is confidential, and the stressed [ə] sound tells you that the [əns] suffix is -ence).

• Nouns that end in -ence or -ance very often have partner adjectives that end in either -ent or -ant – as in confidence and confident. If a noun ends in -ence its partner adjective will always end in -ent. If a noun ends in -ance, its partner adjective will end in -ant: – as in defiance and defiant.

Angel, angle. With these two the main question is where to put that <e>, and the answer is straightforward: In angel that <e> has to come right after the <g> to make it a soft <g>, [æŋˈgɛl], but in angle the <g> is hard, [æŋˈgɛl], so the <e> cannot come right after it.

Angles, Saxons, Jutes, and Frisians were Germanic tribes from northern Europe that invaded Britain in the 5th and 6th centuries. They were all from the Jutland peninsula, roughly modern Denmark, and nearby coastal lands and islands. The Jutes were from northernmost Jutland, the Angles from just south of them. The Saxons and Frisians were south of the Angles, with the Frisians primarily occupying the low coastal areas of the North Sea, essentially modern Belgium and Holland, while the Saxons were primarily on the coast of the western Baltic Sea and areas south in present day Germany. After the departure of the Romans, the Britons invited the Saxons and others to come as mercenaries, to help fight against attacking tribes from the north and the Irish from the west. Thereafter it became more of a full-fledged migration.

The name Angles is the source of the name England; Angle and Saxon the source of Anglo-Saxon. The language of the Frisians is the language closest to Old English.

Apostrophe. Apostrophes are part of spelling and they have quite specific uses:

• First, most of the times they work with <s> to turn nouns into possessive adjectives, as in Jack’s dog. But there are no apostrophes in possessive pronouns: The dog wagged his tail and the dog wagged its tail. And the same is true of the other possessive pronouns hers, ours, theirs, and yours.

• Second, apostrophes are used in contractions like don’t, we’re, and
couldn't. Notice that in the homonyms its and it's, its is a possessive pronoun with no apostrophe, while it’s is a contraction with one. (A mnemonic: “There’s no apostrophe in the possessives his and its.)

- Third, and quite rarely, apostrophes are used to form plurals of letters and numbers: three c’s, five 4’s.

In Greek and Latin an apostrophe was a turning away, and today we use the word apostrophe to refer to the act of speaking directly to an absent or imaginary listener.

**Artificial, beneficial, beneficiary, deficit, efficient, efficiency, office, officer, official -- and others.** Each of these nine CommonWords contains a base that means “to make, do” – sometimes spelled fic1, sometimes fice1:

- artificial = arti+ficé1+ial)
- beneficial = bene+fic1+ial)
- deficit = (de+fic1+it)2
- efficiency = (ex+f1+fic1+iency)
- efficient = (ex+f+fic1+ient)
- office = of2+fice1
- sacrifice = sacri+fice1
- sufficient = (sub+f+ficé1+ient)

In artificial the base arti means “skill, craft”; in beneficial bene means “good, well”; in office of2 means “work, produce”; and in sacrifice sacri means “sanctify, make sacred.” In deficit the prefix (de- carries the sense “reversal, negation”; (ex- has the older meaning “out, beyond” but in efficient and efficiency it’s an intensifier meaning “thoroughly, beyond the usual”; (sub- has an older meaning of “below, under” but in sufficient and sufficiently it carries the sense “nearly, barely.”

CommonWords also contains the following four related fic1 or fice1 words: beneficiary, efficiency, officer, and official. The larger Lexis database contains 118 words with fic1 or fice1, including aficionado, artifice, benefice, coefficient, deficient, edifice, efficacious, immunodeficiency, officious, officiate, orifice, proficient, superficial.

Notice that in fic1 and fice1 the <cc> spells the alveolar fricative sound [s] before <e, i, y> if there is no vowel following the <e, i, y>, as in office or officer. But it spells the palatal fricative [sh] if there is an unstressed vowel immediately after the <e, i, y>, as in official.

**Assimilation.** Assimilation is a historical process that simplifies
pronunciation by changing one of two adjacent sounds to make it more similar to the other. The spelling often changes to agree with the changed pronunciation. In English, assimilation is particularly common in consonants at the ends of Latin prefixes.

In full assimilation the two sounds merge into one, though the spelling does not always change: The full assimilation is reflected in the changed spelling of affair (ad++f+fair, but though the pronunciation is simplified, there is no change in the spelling of adjust (ad+just pronounced [əˈjʊst]). A prefix can have different patterns of assimilation in different settings. For instance, the bilabial sound [m] and its spelling <m> in the prefix (com- do not change in front of the bilabial sounds [b], [m], and [p]: the prefix and stem join via simple addition – combine, commit, and compel. But the [m] and <m> assimilate fully before [l], [n], and [r]: collect, (com+l1+lect5; connect, (com+n1+nect; correct, com+r1+rect.

And most of the time the sound and spelling in (com- undergo partial assimilation, with [m] and <m> changing only partially, to [n] and <n> as in concept, (com+n1+cept and in condemn, confess, congest, conjure, etc. The partial assimilation of (com- to (con- eases pronunciation by moving the points of articulation closer together for the final sound of the prefix and the first sound of the stem. While the [m] at the end of (com- is a bilabial sound pronounced out at the two lips, the sounds spelled by the letters <c>, <d>, <f>, <g>, <j>, <q>, <s>, <t>, and <v> are all pronounced in places in the mouth closer to where the alveolar sound [n] is pronounced.

Assimilation is still with us: If you listen carefully, the word input is often pronounced not [inˈpʊt] but rather [ɪmˈpʊt], with assimilation of [n] to [m] before [p]. The OED and Webster’s 3rd Unabridged even list the variant spelling imput. (EDLL, OCEL, AES, pp. 177-98)

Assure, ensure, insure. These three words have very similar senses. All have sure as their base – (ad+s1+sure, (en1+sure, (in3+sure – and all carry the older sense of sure: “free from danger or harm.” Insure and ensure are essentially interchangeable, though insure is more common in American English, especially in the business sense. Assure carries the base sure’s more modern sense of “certain, confident.” These three stem words can take on a number of affixes to form verbs, nouns, adjectives, and adverbs – for instance, assurance, assuredly, coinsurance, ensurer, ensuring, insurability, insurable, insurance, insuror, reassurance, reassure, reinsure, underinsured, uninsurable, unreassuringly.

The stem word sure also has several different inflected and derived forms and compounds: cocksure, surefooted, surely, sureness, surest, surety, unsure.

Back formation. Back formation is the formation of a new word by taking away part of an older, longer word. For instance, the recent word lase “to
operate a laser” is a back formation from laser, which is itself an acronym contracting “light amplification by stimulated emission of radiation.” An exact parallel is tase from Taser, a contraction of “Thomas A. Swift’s Electric Rifle,” after Tom Swift, the hero of youth fiction. English actually has many back formations: the on-line OED lists 810, including automate from automated or automation, burgle from burglar, gruntled from disgruntled, and pea from pease.

**Bases.** A word's base is the element that contains its core of content and can have prefixes and suffixes added to it. In the word uncounted the base is count (un+count+ed). Bases like count that can stand free as words are called free bases. Bases that cannot stand free as words, like +fect in affect, confection, defective, effective, etc., are called bound bases. (OCEL, AES, pp. 32-39).

**Believe, belief.** The bases lieve and lief “care, love, faith” are variants of one another. The prefix (be-) is an intensifier. This is the same lief that is the old-fashioned word lief “dear, willing.” Notice other <f> and <v> pairs like half, halves; wolf, wolves, elf, elves; shelf, shelves. In Old English the letter <f> spelled the sound [v] when it came between two voiced sounds, such as vowels, and it spelled [f] everywhere else.

**Beowulf.** Beowulf is the most celebrated literary work of the Old English period, probably written sometime in the 5th or 6th century. It is an epic poem in alliterative verse – notice the alliteration of <g>’s and <d>’s in the first line of the sample below. Blending fiction and fact, the poem tells of the feats of Beowulf, a Geat hero from southern Sweden who comes to the aid of Hrothgar the king of the Danes, and fights and kills the monster Grendel and his dam. Later Beowulf becomes king of the Geats and dies in a fight with a fire dragon. The language of Beowulf is so different from Modern English as to be for modern readers essentially a foreign language. These are the opening lines:

```
Hwæt! Wé Gárdena in géardagum
þéodcyninga þrym gefrúnon
hú ðá æþelingas ellen fremedon.
```

(Listen! We of the Spear-Danes in the days of yore,
of those clan-kings heard of their glory
how those nobles performed courageous deeds.)

**Bilabial sounds.** Bilabials are consonant sounds articulated by pressing the two lips together – [b, p, m], as in bob, pop, mom.

**Blends.** Blends are strings of two or three consonant letters that spell two
or more consonant sounds in the same syllable. Some blends are word-initial, more are word-final, a few are both:

- `<bl>` = [bl] as in blue
- `<chr>` = [kr] as in chronic
- `<cl>` = [kl] as in clue
- `<cr>` = [kr] as in crew
- `<ct>` = [kt] as in act
- `<dr>` = [dr] as in draw
- `<fr>` = [fr] as in from
- `<ft>` = [ft] as in soft
- `<gl>` = [gl] as in gloom
- `<gr>` = [gr] as in groom
- `<ld>` = [ld] as in sold
- `<lf>` = [lf] as in shelf
- `<lt>` = [lt] as in belt
- `<mp>` = [mp] as in camp
- `<nch>` = [nch] as in branch
- `<nd>` = [nd] as in brand
- `<nk>` = [nk] as in sank

- `<nt>` = [nt] as in sent
- `<nth>` = [nth] as in tenth
- `<pt>` = [pt] as in slept
- `<qu>` = [kw] as in quarter
- `<rf>` = [rf] as in from
- `<rk>` = [rk] as in mark
- `<rl>` = [rl] as in girl
- `<rm>` = [rm] as in arm
- `<rn>` = [rn] as in barn
- `<rt>` = [rt] as in short
- `<st>` = [st] as in last
- `<str>` = [str] as in straw
- `<sw>` = [sw] as in swell
- `<th>` = [th] as in throw
- `<tr>` = [tr] as in true
- `<tw>` = [tw] as in twin

- `<sk>` = [sk] as in ask
- `<sl>` = [sl] as in sled
- `<sm>` = [sm] as in small
- `<sn>` = [sn] as in sneak
- `<sp>` = [sp] as in spoke
- `<spr>` = [spr] as in spring
- `<squ>` = [skw] as in squeak
- `<st>` = [st] as in last

**Braille.** Braille is a system for converting written text into a tactile form for readers and writers who are partially or completely blind. It is named after its French creator, Louis Braille, who produced his first version of the system at age 15. It consists of raised dots embossed on paper sheets, though today users can also read and write Braille directly from and to their computer screens.

Braille consists of dots that are organized in two-by-three blocks. In the simplified sample below each block of dots represents a letter or **digraph**, a very common word or **suffix**, or a formatting mark:
Another hundred or so blocks represent numerals, punctuation, more formatting marks, and more letter combinations and contractions.

In time Braille will almost certainly be replaced by text-to-speech programs that can ‘read’ written texts and synthesize them into speech, thus allowing visually impaired people to listen to text on a computer.

**British and American spelling.** What we call British spelling was essentially defined by Samuel Johnson in his *A Dictionary of the English Language* (1755), and American spelling was essentially defined by Noah Webster in his *An American Dictionary of the English Language* (1828). Johnson looked to tradition and etymology for his judgements of correct spelling; Webster looked more to his sense of pattern and economy. Three differences between the two spelling systems often cited are (1) the British -*our* vs. the American -*or*, as in *humour*, *humor*; (2) the British -*re* vs. American -*er*, as in *centre*, *center*; and (3) the British tendency to twin final unstressed consonants that American leaves untwinned, as in *traveller*, *traveler*. However, there is considerable lack of agreement among dictionaries, experts, and general readers and writers whether these three distinctions – and all the others often cited in word lists – are strictly British or American. The British and American spelling systems have evolved and coevolved over the centuries to the point that the distinction between them is fuzzy and elusive. For more on this distinction, see my “The Evolution of British and American Spelling” in *The Routledge Handbook of the English Writing System* (V. Cook and D. Ryan, eds.) London & New York: Routledge, 2016. pp. 275-292).

**Capital and lowercase letters.** Capital letters are called capital in the sense of “chief, most important.” Lowercase letters are so called because back in the days of lead type, type setters in print shops stored them in the lower case or rack, as compared with the uppercase, or capital, letters. The use of capital letters in English is quite complicated; the index to the 13th edition of the *Chicago Manual of Style* lists over three dozen references to it. What follows is a simplification. Capital letters are used at the beginning of:

- Sentences: *The rain was pouring down.*
- Proper names: *Lou Gehrig, New York, Grand Canyon, Broadway.*
- Names of deities: *God, Jesus, Jehovah, Allah, Buddha, Zeus.*
- Lines of verse, if the poet capitalized them:

  *Just as my fingers on these keys*  
  *Make music, so the selfsame sounds*  
  *On my spirit make a music, too.*
• Most words in the titles of articles and books: *Pride and Prejudice*.
• The first person singular pronoun *I*.
• *President* when it refers to the President of the United States.

## Capital, capitol

The base in both of these words is *capit* “head, chief.” In *capital* the suffix -al forms adjectives which are often converted to serve as nouns, so the adjective *capital* “main, important” led to the noun *capital* “money.” *Capitol* “the building of the legislature” is from *Capitoline*, the highest of the seven hills of ancient Rome, on which the temple of Jupiter was located. Unlike *capital, capitol* is not used as an adjective. A useful mnemonic: “The capitol has a dome,” with the *<o>* in *dome* helping with the *<o>* in *capitol*.

## Cardinal and ordinal numbers

Cardinal numbers are the ones with which we count: 1, 2, 3, 4, etc. – or *one, two, three, four*. Ordinal numbers indicate places in a numerical order: 1st, 2nd, 3rd, 4th, etc. – or *first, second, third, fourth*. *Cardinal* comes from a Latin word that meant “principal, most important.” *Ordinal* is related to the word *order*.

## Case

Different cases show how a given word relates to other words in a sentence. Modern English has a quite simple case system, relying more on word order and prepositions to show these relationships. In English only nouns and pronouns have different cases. Except for adding *'-s*) to indicate the possessive, or genitive, nouns do not change form to show case. The changes in form for pronouns are more complicated. Unchanged nouns can serve these four cases:

1. the **nominative**, or subjective – that is, the case of the subject of the verb, as in *My aunt sang my uncle a song* – and
2. (2 and 3) the two **objective** cases – that is, the cases of the two objects of the verb: The **accusative**, the case of the direct object – the direct recipient or product of the action of the verb: *My aunt sang my uncle a song*; and the **dative**, the case of the indirect object – the person or thing indirectly affected by the action of the verb: *My aunt sang my uncle a song*. Direct objects answer the question “What?” of the verb: Sang what? Sang a song. Indirect objects answer “To whom or what?” Sang to whom? Sang to my uncle.
3. (4) the **possessive** or genitive case, which, broadly, shows possession and is usually marked with the addition of the suffix *'-s*), as in *My aunt’s voice* or *periphrastically* – that is, by the use of a function word like the preposition *of* in *The voice of my aunt*. 
The changes in form to show pronoun cases are more complicated:

<table>
<thead>
<tr>
<th>Nominative</th>
<th>Accusative</th>
<th>Dative</th>
<th>Genitive</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>I</em></td>
<td><em>me</em></td>
<td><em>me</em></td>
<td><em>my</em></td>
</tr>
<tr>
<td><em>you</em> (singular)</td>
<td><em>you</em> (singular)</td>
<td><em>you</em> (singular)</td>
<td><em>your</em> (singular)</td>
</tr>
<tr>
<td><em>he, she, it</em></td>
<td><em>him, her, it</em></td>
<td><em>him, her, it</em></td>
<td><em>his, her, its</em></td>
</tr>
<tr>
<td><em>we</em></td>
<td><em>us</em></td>
<td><em>us</em></td>
<td><em>our</em></td>
</tr>
<tr>
<td><em>you</em> (plural)</td>
<td><em>you</em> (plural)</td>
<td><em>you</em> (plural)</td>
<td><em>your</em> (plural)</td>
</tr>
<tr>
<td><em>they</em></td>
<td><em>them</em></td>
<td><em>them</em></td>
<td><em>their</em></td>
</tr>
<tr>
<td><em>who</em></td>
<td><em>whom</em></td>
<td><em>whom</em></td>
<td><em>whose, of whom</em></td>
</tr>
<tr>
<td><em>which</em></td>
<td><em>which</em></td>
<td><em>which</em></td>
<td><em>whose, of which</em></td>
</tr>
<tr>
<td><em>that</em></td>
<td><em>that</em></td>
<td><em>that</em></td>
<td><em>of that</em></td>
</tr>
<tr>
<td><em>those</em></td>
<td><em>those</em></td>
<td><em>those</em></td>
<td><em>of those</em></td>
</tr>
</tbody>
</table>

**Caxton, William.** In 1476 William Caxton established the first printing press in Britain and produced his first book, an edition of *Chaucer*’s *The Canterbury Tales*. During the 15 or so years of his career, he printed over a hundred books, mostly works in foreign languages that he and others translated into English. His work was part of the development of Standard Written English and the growing separation between the spoken and written word. It is likely that his tendency to use French words in his translations added to the heavy influence of French on Early Modern English. See French influence.

**Changes in Some Indo-European Sounds.** As the speakers of *Proto-Indo-European* (PIE) spread to the east, west, north, and south, they encountered, settled among, and intermarried with speakers of other languages. This mixing of languages and separation from other PIE speakers changed the languages they spoke. Thus the Indo-European family of languages was born and over the millennia has evolved.

There are several patterned sound and spelling differences among IE languages, which was evidence crucial to the reconstruction of PIE. A very few of these changes are highlighted in yellow in the map below. All of the changes shown here were in the *Germanic* branch of the family, which includes English:
The results of these changes are exemplified below, concentrating on initial stops and comparing Latin with English:

PIE b → Germanic p, thus Latin bullire but English poach “cook in liquid”

PIE d → Germanic t, thus Latin decem but English ten

PIE g → Germanic k, c, thus Latin gelidus but English cold with hard <c>

PIE p → Germanic f, thus Latin ped but English foot
In short, voiced PIE stops all became their unvoiced counterparts in Germanic, while unvoiced PIE stops all became unvoiced fricatives. It was the recognition of patterns like these that, many times over, enabled scholars over the centuries to reconstruct much of how the PIE language must have sounded.

**Changing <y> to <i> and <i> to <y>.** You may be familiar with the old jingle that ends “change the <y> to <i> and add <es>.” But the procedure is broader than that: When we add any suffix that starts with a vowel to a stem that ends in the letter <y> immediately following a consonant, we change the <y> to <i> and add the suffix – for instance, merry + er)02 = meriy+i2+er)02 = merrier.

The replacement does not take place if the <y> is preceded by a vowel because in such cases the <y> is part of a vowel digraph, and digraphs are routinely exempted from such rules: buy + er)01 = buy+er)01 = buyer.

One complication: We add a suffix that starts with an <i> to a stem that ends in a <y> via simple addition: try+ing)1 = trying. But when the suffix starts with any other vowel, and the <y> has a consonant right in front of it, we again change the <y> to <i>: try+i+ed)1 = tried.

Just the reverse of this procedure is also true: When we add a suffix that starts with the letter <i> to a stem that ends <ie>, we change the <i> to <y> and delete the <e>: die plus -ing) = dié1+y4+ing)1 = dying. But when we add a suffix that starts with any other vowel, we just delete the final <e>: die + ed) = dié1+ed)1 = died.

**Chaucer, Geoffrey.** Geoffrey Chaucer (1340?-1400) was the greatest poet of the Middle English period. His work has had an immense influence on the development of English poetry – especially his use of vernacular English and of various poetic forms and rhythms, some of which he developed. His most famous work is the Canterbury Tales, a series of short verse stories told by a group of pilgrims traveling to the cathedral in Canterbury “the hooly blisful martir for to seke.” His language is more accessible to modern readers than is Old English due to the evolution of the language during the intervening centuries. The “General Prologue” opens:

Whan that Aprill with his shoures soote
The droghte of March hath perced to the roote,
And bathed every veyne in swich licour
Of which vertu engendred is the flour . . . .
When that April with his sweet showers
   The drought of March has pierced to the root
   And bathed each vein in such liquor
   Of whose power engendered is the flower . . . .

Clauses and Sentences. A clause is a string of words that contains a subject and a predicate. Clauses come in a somewhat bewildering array of types, but in the most basic sense we can say that the subject of a clause is a noun phrase that refers to the actor – the doer of an action – and the predicate contains a verb phrase that refers to that action. Thus in the clause “The neighbor's dog barked all night” the subject is the noun phrase “the neighbor's dog” built around the noun dog, and the predicate is the verb phrase “barked all night” built around the verb barked. A main, or independent, clause can stand free as a sentence; a subordinate, or dependent clause cannot. Subordinate clauses acting as adverbs are usually introduced by a subordinating conjunction: “because he wanted in out of the rain.” Subordinate clauses can also act as nouns (“He said that he wanted to go home”) or as adjectives (“The man who is sitting in the back row is her father.”)

Clauses can combine to form several kinds of sentences:

- A single main clause forms a simple sentence: “The neighbor's dog barked all night”
- Two or more main clauses form a compound sentence: “The neighbor's dog barked all night, and I didn't sleep a wink.”
- A main clause and a subordinate clause form a complex sentence: “The neighbor's dog barked all night because he wanted in out of the rain.”
- Two or more main clauses and one or more subordinate clauses form a compound-complex sentence: “The neighbor's dog barked all night because he wanted in out of the rain and I didn't sleep a wink.”

Closed syllables. Closed syllables end with a consonant sound. See Syllables.

Code and Performance. Code is the slowly changing system of abstract categories, distinctions, and relationships that structure the English language – including its spelling. It is the current set of rules governing the game of English. Performance is the game itself, as people use code in concrete and particular acts of speaking and listening, writing and reading. Actually code exists in two ways: First, it is the universal, highly abstract, and as yet rather mysterious, set of rules and patterns that inform our language. Second, in addition to this highly abstract and objective code, there are the particular and subjective versions of the code in the minds of the individual users of English. These individual codes reflect the amount of the universal abstract code that individuals have been able to internalize.
from their particular experiences with the language, its users, and the world. The interaction between the abstract and objective code and the concrete and subjective individual codes in acts of performance drives the evolution of the language. For more on “Code and Performance” see Orthography as an Evolving Complex System on this website.

**Comma, coma.** The base of *comma* in *comm*2 “beat, strike, piece cut off, short piece of a sentence.” The base of *coma* is *com*3 “tired, deep sleep.” *Coma* and *comma* are a good instance of the VCV vs. VCCV contrast.

**Compliment, complement.** *Compliment* “praise” is related to *compliant*, “to act according to another’s wish”, which used to mean “polite” – a reminder of that <i> in *compliment*. *Complement* means “to complete”, a reminder of that first <e>.

**Compound words.** A compound word is a word that contains two or more stems. The most easily recognized compounds contain two or more words, or free stems, such as *popcorn*, *pop+corn*; *blackbird*, *black+bird*; *congressman*, *(con+h+gress+man*; *notwithstanding*, *not+with+stand+ing*). This type of compound occurs in three spellings: closed (*firecracker*, *fireman*), open (*fire station*, *firing squad*), and, less commonly, hyphenated (*fire-eater* and the adjective *fire-and-brimstone*). There are several generalizations for suggesting when to choose which one of these three spellings, but so far none are very useful, and the best advice is to check a reputable dictionary.

A second type of compound contains one or more bound bases plus one or more free stems: *biofeedback*, *bio2+feed+back1*; *electromagnet*, *electro+magnet*; *hydroelectric*, *hydro+electric*; *gasohol*, *gas1+ohol*; *ionosphere*, *iono+sphere*; *monoxide*, *mon3+oxide*; *petrochemical*, *petro+chemical*; *thermodynamics*, *thermo+dynamics*. Special cases are compounds that contain a free base plus a contraction: *can’t*, ‘twas, he’d, she’ll.

A third type contains two or more bound bases and no free base: *moped*, *mo2+ped6*; *equilibrium*, *equi1+libr1+ium2*; *centrifuge*, *centri+fuge*; *hydrogen*, *hydro1+gen6*; *isotope*, *iso+tope1*; *litmus*, *lit3+mus2*; *positron*, *posi2+tron1*; *telemetry*, *tele1+metr1+y3*; *gastroenterologist*, *gastro+enter2+olog+ist1*. Such compounds with bound bases are most common in the technical and scientific registers.

**Concatenation.** *Concatenation* means a linking together. It comes from Latin, with the earlier meanings “bind, chain.” In orthography concatenation has to do with how sounds and letters link together, as into blends, doublets, doublet equivalents, digraphs, trigraphs, and simplifications, and with things that occur when elements concatenate: simple addition, twinning, silent final <e> deletion, assimilation, changing <y> to <i> and <i> to <y>. 
**Concatenations.** Some **concatenations** of sounds and letters can pose special problems for spellers. They can look weird and be hard to pronounce, so sounds and letters in the middle of them can be easily lost – or at least misplaced. This is a case where **explication** can help by unfolding the word. The following are some examples of more complex concatenations and how they can be unfolded via explication:

**<tchb>:** Notice that <tch> tends to follow a short vowel:

- *hatchback*    hatch1 + back1
- *matchbook*    match2 + book
- *matchbox*     match2 + box1
- *pitchblende*  pitch2 + blende
- *sketchbook*   sketch + book
- *switchback*   switch + back1
- *switchblade*  switch + blade
- *switchboard*  switch + board
- *watchband*    watch + band1

**<ghtw>:** Notice that <gh> and <ght> tend to follow a long vowel:

- *flightworthy* flight + wor1+th)2+y)1
- *lightweight*  light2 + weigh+t)2
- *nightwear*    night + wear1
- *rightward*    right + ward]
- *straightway*  straight1 + way1
- *tightwad*     tight + wad

**<ndst>:** Notice how that [d] tends to get lost.

- *bandstand*   band2 + stand
- *grandstanding* grand + stand+ing)1
- *grindstone*  grind + stone
- *handstand*   hand + stand
- *sandstone*   sand1 + stone
- *sandstorm*   sand1 + storm
- *soundstage*  sound1 + stage
- *standstill*  stand + still1
- *windstorm*   wind1 + storm

**Concrete vs. Abstract, and Specific vs. General.** People often treat **concrete** and **specific** as **synonyms**, as they do with **abstract** and **general**. But **concrete** means perceptible by the senses – by sight, sound, smell, taste, touch; and by contrast **abstract** means not perceptible by the senses, immaterial, conceptual. **Specific** means dealing with a single or a very limited set of things; and by contrast **general** means dealing with a larger, even infinite, set.

If we treat the concrete-abstract contrast as a vertical axis, and specific-general as a horizontal, we can identify a cognitive space with four realms:
Concrete-specific is the realm of experience and its particulars; abstract-specific is the realm of logic and measurement. Concrete-general is the realm of analogy and metaphor; abstract-general is the realm of universals.

We can also identify lines of thought through this cognitive space with some representative items and modes of thinking in each realm:

The philosopher Karl Popper argues that human reality consists of three worlds – Worlds 1, 2, and 3. His World 1 is the physical world – the outer world of physical objects and their events. It provides the substrate for our concrete-specific cognitive realm. World 2 is the inner world of psychological states and responses, of feelings, volitions, perceptions, the world within the individual human mind. World 2 is represented by our concrete-specific, concrete-general, and abstract-specific realms. World 3 is the symbolic world of abstract general values, principles, and laws, the product of the human mind. (See, for instance, Popper’s *In Search of a Better World: Lectures and Essays from Thirty Years*. London and New York: Routledge, 1984, 1992, pp. 7-26.)

**Confident, confidant.** Both *confident* and *confidant* come from a Latin word that meant “to rely on.” The adjective *confident* is discussed at -ance), -ence); -ant), -ent). The noun *confidant* is pronounced [cõnˌfi·dánt’], with a short <a> in <dant> where *confident* has a schwa in <dent> and with a different stress pattern. *Confidant* is often pronounced with an [ä] rather than an [ã], and is spelled *confidante* when referring to a woman or girl.
Conjunctions. In grammar there are two types of conjunctions, or joiners: coordinating and subordinating. **Coordinating conjunctions** join together two things that are being treated as equals, or coordinate – for instance, the adjectives blue and white in the phrase “the blue and white orchids,” or the two nouns in the phrase “dogs and cats,” or the two main clauses in the compound sentence “She said she’d be here and she usually keeps her word.” Other coordinating conjunctions are but, or, so, and nor.

**Subordinating conjunctions** combine two unequal clauses and mark the subordinate clause, whether it comes before or after the main clause: “If it rains, we’re not going” (with a comma between the clauses) and “We’re not going if it rains” – usually with no comma. There are many subordinating conjunctions but common ones are after, although, as, because, before, if, since, that, though, unless, until, while. Subordinate clauses usually cannot stand free as sentences. The relation between main and subordinate clauses is much like that between free and bound stems or free and bound bases.

Conscious, conscience. The base of these two words is sci1, which originally meant “to cut or split” and acquired the later sense of “to separate one thing from the other” and finally in Latin came to mean “knowledge, understanding.” Sci1 also occurs in the words science, scientist, scientific, omniscient, etc. The suffix in conscious is -ous), which forms adjectives that can be converted to nouns. The suffix in conscience is -ence), which forms nouns but not adjectives. Paying special attention to the pronunciation of the end of these two can help. And a possible mnemonic: *Conscience* has two <n>’s, just like noun.

Consonant letters and the sounds they spell. These are the 22 consonant letters:

- <b> <c> <d> <f> <g> <h> <j> <k> <l> <m> <n>
- <p> <q> <r> <s> <t> <u> <v> <w> <x> <y> <z>

<b> and [b]. The letter <b> and its doublet <bb> regularly spell the sound [b], as in *bib* and babble, accounting for 99.9% of the spellings of [b] in CommonWords. The <bb> doublet is rarely due to simple addition (dumbbell, subbranch), sometimes due to assimilation (abbreviate), is more often due to twinning (chubby, dubbing, flabby, jobber, robbery, rubber, shrubbery, stubby), or to the VCV/VCCV contrast (abbey, cabbage, hobby, jabber (“talk”), rabbit, ribbon, rubbish) or the VClE/VCCle contrast (bubble, pebble, rabble). The sound [b] is also spelled with the simplifications <mb> (bomb, comb, lamb, etc.) and <pb> in cupboard. AES, pp. 328-33.

<c> and [k] and [s]. For <c>=[k] see <k> and [k]. The letter <c> spells [s] only when it’s immediately followed by <e>, <i>, or <y> – as in ice, icing, icy. In CommonWords [s] is spelled <c> about 20% of the time;
<s> about 75% of the time. See also Soft and hard <c> and <g>. AES, pp. 405-06.

<d> and [d]. In CommonWords the sound [d] is spelled <d> or its doublet <dd> about 99.5% of the time, as in dad and riddle. The <dd> doublet is due to assimilation only in sudden, sometimes due to simple addition (addict, address, adduce, midday), more often due to twinning (bidden, daddy, goddess, muddy, reddest, rittance, sadden, wedding), and to the VCV/VCCV contrast (bladder, eddy, ladder, shudder, udder) and the VCle/VCCle contrast (fiddle, huddle, middle, paddle, riddle, saddle). The sound [d] is also spelled with the simplification <ld> in could, would, should, and solder. And the letter <d> also is used in the two doublet equivalents <dg> and <dj> to spell the sound [j], as in bridge, edge, judge, knowledge, and adjective, adjourn, adjourn, adjust. When we add the past suffix -ed to any verb that ends with a voiced sound other than [d], the <ed> spells the sound [d]: boomed, filed, freed, starred, sighed, sinned. AES, pp. 337-42.

<f> and [f]. In CommonWords about 86% of the time the letter <f> and its doublet <ff> spell the sound [f], as in fluff. But <f> is very rare at the end of words after a vowel: The only known cases are if and of, and in of <f> spells [v], the only known instance of <f> = [v]. Usually word-final [f] after a vowel is spelled <ff>: cliff, fluff, gruff, off, scoff, sniff, staff, stuff, whiff. The <ff> doublet is also often due to simple addition (griefful, safflower, shelfful, soffit, traffic), more often due to assimilation (affection, afford, differ, difficult, effect, effort, offend, offer, office, suffocate), rarely to twinning (iffier, iffy, reffed, reffing), more often to the VCV/VCCV contrast (buffalo, caffeine, coffee, daffodil, muffin, off, tiffany) and the VCle/VCCle contrast (ruffle, waffle). The sound [f] is also spelled with the simplifications <ft> and <lf>, as in often, soften and calf, behalf, half. The sound [f] is spelled <gh> in cough, enough, laugh, laughter, rough, tough. And it is spelled <ph> and <pph> in phone, photograph, prophet, sphere, triumph, typhoon and sapphire. For <gh> and <ph> see Digraphs <ch, gh, ph, rh, sh, th, wh>. AES, pp. 377-84.

<g> and [g]. The letter <g> accounts for 99% of the spellings of [g] in CommonWords and for 67% of the spellings of [j]. See Soft and hard <c> and <g>. The letter <g> regularly spells the sound [g] before the vowel letters <a, o, u> or any consonants, as in gate, glove, got, great, gut. Sometimes in the <gu> spelling the <u> is a silent insulator used to separate the <g> from an <e, i, y>, which would spell the sound [j] rather than [g]: disguise, fatigue, guess, guide, guilt, guy, vague. The <gg> doublet is due to simple addition (doggone, only known case), assimilation (aggrandize, aggravation, aggregation, aggrieve), twinning (baggy, beggar, druggist, logging, luggage, wagged), or the VCV/VCCV contrast (dagger, maggot, toboggan, trigger), or the VCle/VCCle string (goggles, squiggle, struggle, wiggle). The digraph <gh> spells [g] in ghastly, ghoul, ghost,
sorghum, and spaghetti. Due to simplification the sound [g] is spelled <tg> in mortgage mort1+gage1 and <ckg> in blackguard black+guard. AES, pp. 350-55.

<h> and [h]. In CommonWords about 97% of the time, the sound [h] is spelled <h>, as in hat, house, hungry. The letter <h> also occurs in these digraphs and trigraphs:

- <ah> to spell [a] and [ä] as in Allah, hurrah, Jehovah, messiah
- <aoh> to spell [ō] as in pharaoh
- <ch> to spell [ch] as in church and [k] as in school and stomach;
- <sh> to spell [sh] as in shush;
- <th> to spell [th] as in thin, [th] as in then, and [t] as in Thomas;
- <rh> to spell [r], as in rhinoceros, rhubarb, rhythm;
- <rrh> to spell [r], as in diarrhea, hemorrhage, myrrh;
- <wh> to spell [h], as in who, whole, whom, whose, or to spell [ˈw] or [w], as in whale, while, whether;

The letter <h> occurs as a true silent letter in heir, honest, honor, hour and after <x>, as in exhaust, exhibit, exhilarate, exhort, exhume. AES, pp. 386-90.

<j> and [j]. The letter <j> regularly spells the sound [j], nearly always at the head of bases, as in, joke, junior, project, rejoice, but accounting for only 24% of the spellings of [j] in CommonWords. The sound [j] is spelled <g> before <e, i, y> – as in cage, caging, cagy, accounting for 67% of the spellings of [j]. See Soft and hard <c> and <g>. The sound [j] is spelled with the doublet <gg> only in exaggerate. After short vowels it is often spelled with the doublet equivalent <dg> as in abridgment, badge, fledgling, judgment, ridge, lodger, fudging. It is less often spelled with the doublet equivalent <dj> as in adjective, adjoin, adjourn, adjust. And it is spelled with a palatalized <d> usually followed by <u>, as in fraudulent, gradual, graduate, individual, procedure, schedule. The letter <j> spells [y] in the Norwegian adoption fjord. It spells [h] in the Spanish adoption junta, and due to simplification <ju> spells [w] in marijuana. AES, pp. 417-21.

<k> and [k]. The sound [k] has the most complicated spellings of all the consonants. The letter <k> regularly spells [k], but that is just the beginning, and all that follows is still a simplification. The complications arise because the sound [k] has five major spellings (<c>, <ch>, <k>, <q>, <qu>) and four doublets or doublet equivalent spellings (<cc>, <ck>, <cq>, <kk>). The numbers in parentheses below show the number of words in CommonWords with each spelling. The sound [k] is spelled

- <c> before <a, o, u>, as in cat, cot, cut (1412)
- <ch> as in chaos, chemistry, chlorophyll, echo, schedule, school, stomach – and ache (87)
- <k> before <e,i,y>, as in shake, shaking, shaky (309)
- <q> as in equal, frequent, quack, quarrel, queen, request, squeak, squirrel (91)
- <qu> as in antique, clique, croquet, grotesque, mosque, mosquito, opaque, queue (10)

The sound [k] is also spelled
- <cc> due to assimilation, as in account, accurate, occasion, occur, and due to twinning in sicced (31)
- <ck> in a VCC string, as in back, chicken, lucky, package, tackle (118)
- <cq> due to assimilation, as in acquaint, acquire, acquittal (4)
- <kk> rarely, due to twinning, as in trekked, trekkie, yakking (0)

Other minor spellings include <cch> in saccharine, zucchini (0); <kh> in khaki (0); <lk> in talk, walk (10); <sc> in viscount (0); <x> in excel, excite (13). AES, pp. 355-72.

- <l> and [l]. The letter <l> regularly spells the sound [l], and the sound [l] is spelled either <l> or <ll> over 99% of the time in CommonWords. The contrast between the <l> and <ll> spellings is less clear-cut than it is for most other consonants. But the <ll> doublet is often due to simple addition (actually, soulless), assimilation (allege, alloy, collapse, collect, illogical, illusion), twinning (compelling, controller, patrolling, propellant, rebellion), or the VCV/VCCV contrast (ballad, collar, lullaby, silly, valley, yellow). The only other spelling of [l] is the simplification <sl> in aisle, isle, island. AES, pp. 439-47.

- <m> and [m]. The letter <m> regularly spells the sound [m], and the sound [m] is nearly always spelled <m> or <mm>, more than 98% of the time in CommonWords. The <mm> doublet is always due to simple addition (command, commercial, commit, communal), assimilation (immature, immediate, immortal, symmetry), twinning (dammed, dimmest, drummer, humming, slimmer), or the VCV/VCCV contrast (comma, dilemma, hammer, mummy, simmer). There are four minor spellings of [m], all simplifications and usually at the end of words:
  - <mb> as in bomb, climb, comb, dumb, lamb, tomb, womb;
  - <lm> as in the most common pronunciation of alms, calm, palm, psalm, qualm;
  - <mn> as in autumn, column, condemn, damn, hymn, solemn;
  - <gm> as in diaphragm, paradigm, phlegm, syntagm. AES, pp. 423-29.

- <n> and [n]. The letter <n> regularly spells the sound [n], and about 99% of the time the sound [n] is spelled either <n> or <nn>. As usual the <nn> doublet is always due to simple addition (innate, innocent, innuendo, leanness, thinness), assimilation (announce, connect, connotation), twinning (beginning, funny, inner, thinnest), or the VCV/VCCV contrast (cannibal, channel, minnow, penny, tennis, tunnel). In four simplifications the sound [n] is also spelled
• <kn> as in knead, knee, knew, knife, knight, knot, knuckle;
• <gn> as in gnarl, gnash, gnat, gnaw, gnome;
• <pn> as in pneumatic, pneumonia, pneumococcus;
• <mn> as in mnemonic. AES, pp. 429-35.

<p> and [p]. The letter <p> and its doublet <pp> regularly spell the sound [p], more than 99% of time in CommonWords – <p> at any position in a word, <pp> never at the beginning or end. The <pp> doublet is rarely due to simple addition (one pronunciation of lamppost, stepparent), often due to assimilation (appeal, approach, approve, oppress, supply, support, suppose), more often due to twinning (clipping, happy, slipper, upper, zipper), the VCV/VCCV contrast (copper, hippopotamus, kipper, pepper, poppy) and the VCle/VCCle contrast (apple, cripple, grapple, ripple, supple, topple). AES, pp. 335-37.

<q> and [k]. See <k> and [k].

<r> and [r]. The letter <r> and its doublet <rr> regularly spell the sound [r], and account for more than 98% of the spellings of [r] in CommonWords. But the consonant sounds [l] and [r] are classified as liquids, and with both liquids the VCV/VCCV distinction is less useful than with other consonants – consider, for instance, the effect of [r] on a preceding vowel, as in fate or fame or faze vs. fare. However, <rr> spells [r] in any position in the word: acre, cheer, circle, direction, miracle, orange, player, rabbit, ruin, theory. The doublet <rr> sometimes spells [r] at the end of words: bizarre, burr, err. Often the <rr> is due to twinning: deterrent, furry, occurred, recurrent, starry, stirring, warrior. Fairly often it is due to assimilation: arrive, correct, corrupt, irradiate, irresponsible, irrigate, surrogate or to simple addition: earring, interrogate, interrupt, surreal, surrender, surround. The <rr> also can occur in a VCC string, though it can be hard to identify a short vowel at the head of it: current, horrible, marry, merry, worry, sorry, squirrel, stirrup, warrior. Three minor spellings of [r] are the simplification <wr> (wrap, wreck, wrench, wrinkle, write, wry), and the digraph and trigraph <rh> and <rrh> (rheostat, rhinoceros, rhubarb, rhythm; diarrhea, hemorrhage, myrrh). In corps [kôr] we apparently must say that [r] is spelled <rps> via simpification. AES, pp. 447-55.

<s> and [s]. In CommonWords 83% of time [s] is spelled <s> or <ss>, as in sassy and sissy. The letter <s> spells [s] anywhere in a word: solid, dense, radius. The doublet <ss> is rarely due to twinning (busses, gassed, plusses, yesses, sissy), sometimes due to simple addition (dissatisfied, dissect, dissipated, misspeak), often due to assimilation (assault, assembly, assets, assignment, assimilation, assist, assorted), sometimes due to the VCCV/VCV contrast (blossom, casserole, opossum pessimism, vessel), and rarely due to the VCCle/VCle contrast (hassle, tussle).
However, things get much more complicated: First, the sound [s] is often spelled <c> before <e, i, y>: lace, lacing, lacy. See Soft and hard <c> and <g>.

A second complication is that very often the letter <s> spells the sound [z], more often than does the letter <z> (capitalism, dogs, has, observe, visit, dismal), while <ss> also sometimes spells [z], as in (dessert, dissolve, possessive, scissors).

A third complication is that <s> appears in several simplifications that spell [s]: <ps> (psychic, psychology), <sc> (scene, scientist), <st> (listen, whistle), <sh> (isthmus), <sw> (sword).

And finally <s> appears in spellings of sounds other than [s]: <tsch> = [ch] in putsch, <sl> = [l] in aisle, isle, island, and four spellings of [sh]: <s> in dimension, insurance, sugar, surely; <sc> in conscience, conscious, crescendo, fascism; <sh> in short, bishop, blush; <ss> in issue, mission, passion, session, tissue.

In a word-final base that would otherwise end with a single <s>, if there is a consonant or vowel digraph preceding the <s>, a silent final <e> is added as an insulator to keep the word from looking like a plural: browse, curse, dense, else, hearse, lapse, goose, moose, please, sparse – and by extension clause, house, noise, etc. AES, pp. 394-405, 407-12, 422.

<t> and [t]. About 97% of time in CommonWords the sound [t] is spelled <t> or <tt>. The letter <t> spells [t] at the beginning, middle, or end of a word: testament, thermostat, treatment. The doublet <tt> is often due to twinning (admittance, flattest, knitting, rebuttal, rotten, shutter), to simple addition in some pronunciations (cattail, outtake, rattrap), to assimilation (attack, attend, attention, attire, attribute), to the VCC/VCV contrast (attic, butter, cotton, glitter, jetty, lettuce, motto, utter), or to the VCCle/VCle contrast (battle, bottle, brittle, cattle, kettle, scuttle, shuttle, throttle). But after this ruly start there are several unruly spellings of [t] – usually quite rare, sometimes simplifications involving the letter <t>, sometimes Latinizations:

- <ed> or <d>= [t] in the past tense of verbs that end with one of the voiceless sounds [ch, f, k, p, s, sh] (perched, miffed, hooked, stopped, kissed, fished) – depending on whether we treat the <e> as silent, making <d>=[t] (25 instances in CommonWords)
- <ght> = [t] (bought, caught, eight, freight, height, sight, weight) (76)
- <bt>=[t] in some Latinizations (doubt, debt, subtle) (7)
- <pt>=[t] (pterodactyl, ptomaine, receipt) (1)
- <th>=[t] (Thomas, thyme) and one pronunciation of posthumous (2)
- <tw> = [t] in two (1)
- <cht> = [t] in yacht (1)
- <ct> = [t] in the Latinization indict (1)
- <dt> = [t] in the Dutch veldt (0)

In one pronunciation of eighth [ätth] there is apparently an “invisible
"[t]," the reverse of a silent letter. AES, pp. 342-49.

<u> and [w]. The letter <u> usually spells vowel sounds, but it sometimes spells the consonant sound [w], and whether it spells [w] or not, it is a consonant whenever it follows <q>. See <k> and [k]. The letter <u> sometimes spells [w] after <g>: anguish, extinguish, iguana, jaguar, language, linguist, penguin. It rarely spells [w] after <s>: assuage, persuade, suave, suede, suite. It spells [w] after <j> only in marijuana and after <p> only in pueblo. In words like use, utility and uniform <u> spells the complex [yû] sound, which some linguists treat as a consonant-vowel sequence. AES, pp. 342-49.

<v> and [v]. Although it took centuries for <u>, <f>, and <v> to sort themselves out in our alphabet, things are quite simple now for the letter <v>: It and its very rare doublet <vv> regularly spell [v], as in verve and revved. The <v> spelling can occur anywhere in a word although at the end of words <v> is usually covered by the insulator <e>: advise, alive, shove, silver, valve, verdict, wolves. The <vv> doublet is very rare and restricted to new and rather peripheral words: chivvy ("vex, harass"), flivver, divvy, (probably formed from divide), navvy ("laborer," formed from navigator), revving, savvy, skivvies ("underwear"). The two minor spellings of [v] are <lv>, as in calve, halve, salve, and <f>, only in of. AES, pp. 373-77.

<w> and [w]. The letter <w> regularly spells the consonant sound [w], as in award, sweet, swift, swoop, waitress, wisdom, wonder, wooden. A second common spelling of [w] is <wh> in such words as awhile, overwhelm, whale, wheel, when, where, whisper, white, why. Most words with <wh> were spelled <hw> in Old English and pronounced [hw]. The Norman scribes inverted the spelling, bringing it into line with other consonant digraphs ending in <h>: <ch, gh, sh, th>. Today the pronunciation of <wh> varies from [w] through [hw] to [hw]. The sound [w] has an unusual spelling in one, once – where apparently we have to say either that [wû] is spelled <o> or we have an “invisible” [w]. The letter <w> appears as a vowel letter in the digraphs <aw, ew, ow>: maw, mew, mow. See <u> and [w]. AES, pp. 456-59.

<x> and its sounds. At the beginning of words <x> usually spells the sound [z]: xenon, xenophobia, xylophone, xerox. But in words where the <x> is treated as an initial, it spells [ɛks] – as in x-ray and the verb x, as in x’d out, x’ing out. Often, and especially in final position, <x> spells the voiceless [ks]: climax, dexterity, exclaim, express, external, fox, index, next, paradox, textile, wax. In the middle of words it spells [ks]’s voiced counterpart [gz] when it precedes a stressed vowel: exam, exasperate, executive, exempt, exhaust, exhibit, exist, exuberant. In English the doublet <xx> does not occur, probably because most of the time it is a one-letter sound cluster, [ks] or [gz]. See AES, pp. 352, 371.
<y> and [y]. The letter <y> usually spells vowel sounds. But at the beginning of elements and syllables it regularly spells the consonant sound [y]: beyond, farmyard, yacht, yawn, yellow, yesterday, youth, yule. In one pronunciation of llano “grassy plain” [y] is spelled <ll>. AES, pp. 459-60.

<z> and [z]. In spelling the sound [z] <s> is actually more common than <z> (nearly 8:1), and the <s>, <z>, <ss>, and <zz> spellings of [z] can be a bit of a tangle. The sound [z] is spelled in the following ways:

- <z> at the beginning, middle or end of words: breeze, citizen, frenzy, ooze, quiz, sneeze, topaz, zebra, zigzag, zinc, zone, zoo (83 instances in CommonWords)
- <zz> due to twinning: quizzed, whizzing, fezzes
- <zz> in VCC strings: blizzard, buzz, buzzard, dizzy, fuzz, gizzard, jazz;
- <zz> in VCCle strings: drizzle, embezzle, puzzle, sizzle (all 3 <zz> spellings total only 7 instances in CommonWords)
- <s> in the middle or at the end of words when it is preceded or followed by a voiced sound: closet, cosmic, dogs, has, lunches, observe, pose, was, wisdom – and words with the suffix -ism) (capitalism, socialism, etc.) and when the suffixes -es1 and -es2 are pronounced [iz] when added to 3rd person present verbs and nouns that end with [s, z, sh, ch] (621 instances in CommonWords)
- <ss> via simple addition: dessert, dissolve, possess;
- <ss> in a VCC string in scissors. (both <ss>’s add up to only 6)

The sound [z] has these minor spellings: <x> in initial position (xenon, xylophone); in the three variants <cz> in czar, <ts> in tsar, and <tz> in tzar; and <sth> in asthma.

Silent final <e> is sometimes used to insulate otherwise final <s> and <z>, especially after consonants or vowel digraphs: fez, quiz, topaz but bronze, gauze, squeeze, wheeze; and as, lens but applause, disease, noise. See AES, pp. 391-97.

Consonant sounds. English consonant sounds are usually categorized by place and manner of articulation, as in the chart below. Columns indicate places of articulation in the mouth; rows indicate manner of articulation. In the chart below, in cases where a single set of square brackets contains two sounds, the first of the two is voiced, the second unvoiced, or voiceless. See Voiced vs. voiceless (unvoiced) sounds. Characters in diagonals are from the International Phonetic Alphabet:
Bilabials are articulated at the two lips; labiodentals at the lower lip and upper teeth; interdentals with the tip of the tongue between the upper and lower teeth. Alveolars are articulated with the front of the tongue near or touching the alveolar ridge behind the upper teeth. Palatals are articulated with the tongue against the palate, or roof of the mouth; velars are articulated with the back of the tongue at or near the soft palate at the rear of the mouth; glottals are articulated at the glottis further back and above the vocal cords.

Stops are articulated by stopping the flow of air and releasing it suddenly. Some linguists call them plosives. Fricatives are articulated with enough closure to produce friction. Affricates begin with a stop and end with a fricative: \([j] = [dzh]\), \([ch] = [tsh]\). Nasals are articulated through the nose. Nasals, liquids, and semivowels are articulated smoothly, with no friction. *(EDLL, OCEL, AES, pp. 201-06, 209-12)*

**Consonant and vowel letters.** In general we treat a letter as a consonant letter when it spells a **consonant sound** and as a vowel letter when it spells a **vowel sound**. The three letters \(<y, u, w>\) sometimes spell consonant sounds, sometimes vowel sounds, so they can be either consonant or vowel letters:

The letter \(<y>\) is a consonant only at the beginning of syllables when it spells the consonant \([y]\) sound that it spells in *yell*; everywhere else it is a vowel.

The letter \(<u>\) is a vowel when it spells *vowel sounds* such as \([\ddot{u}]\) in *buck*, \([\ddot{u}]\) in *bull*, \([\ddot{u}]\) in *include*, \([y\ddot{u}]\) in *butte*, and in the *vowel digraphs* \(<au>\) as in *author*, \(<eu>\) as in *neutron*, \(<ui>\) as in *build*, \(<ou>\) as in *out*. It is a consonant in only two situations: (i) when it spells the
consonant sound [w], as in language and pueblo; and (ii) when it follows the letter <q>, whether it spells [w], as in quit, or does not, as in mosquito and plaque.

The letter <w> is usually a consonant, including in the consonant digraphs <wh> and <wr>. It is a vowel only in the vowel digraphs <aw>, <ew>, and <ow>, as in craw, crew, and crow. (In some mercifully rare Welsh adoptions it spells long <u>, as in cwm [kûm] “a mountain hollow.”)

The following four are always vowel letters: <a, e, i, o>. And the following nineteen are always consonant letters: <b, c, d, f, g, h, j, k, l, m, n, p, q, r, s, t, v, x, z>. AES, pp. 207-12. For more on consonant and vowel letters see Sometimes a Vowel Is Not a Vowel, and Sometimes a Consonant Is and On Vowels and Consonants – All You Ever Wanted to Know, But . . . in the Short Articles branch of this website.

Content and meaning. Content is what words have; meaning is what people make of that content. In spite of the way we commonly use the words mean and meaning (including in this compendium), words don't actually mean; people do. Words and elements have at least two kinds of content: (i) Semiotic content, an agreed-upon range of senses that people can draw from to formulate and communicate their own meanings. And (ii) morphological or syntactic content, information about the function of the word in sentences. For instance, the suffix -s)2 indicates that the word is a 3rd person singular verb – as in she sings sing1+s)2, and -s)3 indicates that the word fits is a plural noun – as in the fits fit4+s)3.

The study of content is semiotics; the study of meaning is semantics. Although the distinction between content and meaning is very important, it is usually ignored, and meaning is often used when what is actually being referred to is content. AES, p. 35. For more on this distinction see “Content and Meaning” in Orthography as an Evolving Complex System in the Short Articles branch of this website.

Contraction. The most easily recognized contractions are those that contain apostrophes – like doesn’t and o’clock, which contract – or abbreviate – does not and of the clock. In earlier English there was a construction known as the his-possessive – as in Shakespeare’s “And art thou not Poines, his brother?” – meaning “the brother of Poines.” Though the issue is not completely clear, it seems likely that our possessive -‘s)1 suffix contracts that earlier use of his. Or it may simply contract the Middle English possessive suffix -es).

But not all contractions in English spelling are marked with apostrophes: CommonWords lists 60 bases, 28 prefixes, and 13 suffixes as instances of contraction: Among bases wel as in welcome is a
contraction of the free base well “good”; yup as in yuppie contracts young urban professional; and syn as in synfuels contracts synthetic, a kind of contraction particularly common in the scientific and technical registers – for instance, ket1 as in phenylketonuria contracts ketone “a chemical compound.”

**Conversions.** Conversion is the process by which a word of one part of speech is converted for use as a different part of speech. Since English does not use many inflections to indicate parts of speech, conversion is very easy, productive – and ancient – in English. The conversion of nouns to verbs, and vice versa, is probably the most common: CommonWords lists 2,523 words tagged as both regular noun and regular verb: age, jump, kiss, lodge, milk, neglect, outrage, etc. Even more creative conversions are possible – for instance, the conjunction-to-verb-and-noun nonce conversion in “But me no buts.” Odd but understandable. Conversion is one of the things that can make studying parts of speech difficult for students.

Also, CommonWords lists 57 cases where affixes are converted for use as bases. For instance, in the informal ref, which contracts referee (re1+fer+ee)1, the prefix (re1- is converted to part of the new base ref.

**Council, counsel, consul, consult.** These four words have a lot in common: They all contain a prefix meaning “with, together”; (coun- is just a variant of (con-, which is an assimilated form of (com-. The base cil1 in council and the base sel1 in counsel both descend from a root that originally meant “to shout.” Both council and counsel came by way of the Latin word consilium, which meant among other things “a deliberative or advisory body” and “advice.” So neither etymology nor word structure are much help in trying to sort out these two. About all we can say is that usually council is a noun, referring to an advisory group and its advice, while counsel can be a verb meaning to offer advice or a noun referring to the person offering it.

Consul and consult are also tightly bound together: Both contain the assimilated prefix (con- “with, together.” The bases sul and sult2 both came from a root that meant “to take, grasp, gather.” Both words came by way of Latin words that related to the idea of taking council. Consul is a noun referring to the official representative of a nation; consult is almost always a verb “to seek or give advice,” and when it is a noun, it usually involves doctors and medicine.

**Decent, descent, dissent.** The adjective decent explicates to dec2+ent), the noun descent to (de+scent2, and the noun and verb dissent explicates to (dis+sent2. The earliest meaning of dec2 was “take, accept,” which in time became “to be fitting, acceptable.” The base scent2 means “leap, climb,” the prefix (de- means “down, lower.” Descent, a noun, refers to the result of the verb descend. With different prefixes, ascent and ascend and
extent, extend have that same noun/verb relationship.  
In dissent the base sent2 means “head for, go” which is negated by
the prefix (dis-. Notice that unlike decent, both descent and dissent have
stress on the second syllable.

Dental sounds. Dentals are consonant sounds pronounced with the
tongue against or near the upper dental ridge – [d, t, z, s, n, l, r]. Also
called alveolar sounds.

Derivation, derived forms. Derivation is the process by which words are
derived, or formed, by adding suffixes (and less often prefixes) to stems. See Derivational Suffixes.

Desert, dessert, deserts. These three can be quite a tangle. Desert can
be either a noun “a barren place” with stress on the first vowel, or a verb
“to abandon, strand” with stress on the second vowel. As a verb it is a
homophone of dessert “the final course of a meal,” though in meaning and
history they are not at all related. The nouns desserts and deserts – as in
He got his just deserts – are also homophones unrelated in meaning and
history.

  Desert explicates to (de- “off, away” plus the base sert1 “line up,
  attach” – and thus the sense “unattached, abandoned.”

  Dessert explicates to (des2+sert2. The prefix (des2- is a French
  form of (de-. The base sert2 originally meant “slave” but came to mean
  “serve” – so dessert originally meant “the removal of that that was served.”

  Deserts explicates to (de+sert2+s)3 and is related to the word
deserve.

  The major spelling problem is remembering that <ss> in dessert. A
mnemonic can help: “Dessert tonight is strawberry shortcake.”

Diacritic marks. Diacritic (dia+crit+ic)2 comes from a Greek word
meaning “to separate or distinguish.” Diacritic marks in modern English
spelling are all holdovers from their foreign source words, especially
Romance and German, where they served to distinguish various aspects of
pronunciation. In modern English they have usually lost their original
diacritic function, no longer distinguishing sounds. With four limited
exceptions today they serve only an etymological function: identifying the
words in which they occur as noteworthy foreign and usually part of what
might be thought of as the “high class” register.

The four diacritics still used in English to distinguish sounds are these:

  • the acute accent used in Romance words to indicate that the final <e> is
    pronounced, as in abbé, appliqué, attaché, émigré, sauté, touché;
  • the cedilla <¸> used in Romance words like façade, garçon, curaçao,
    and soupçon to indicate a soft <c>;
  • the umlaut or diacritical mark <`> sometimes used to indicate that a vowel is
pronounced, as in Brontë, naïve, noël – or was umlauted in German, as in fräulein;

- the tilde  \( ñ \) used over \(<n>\) in Spanish adoptions to indicate that the \(<n>\) is pronounced \([ny]\), as mañana, piñata, piñon, piraña (also piranha), and señor;

The breve  \(  \), macron  \( ¯ \), overdot  \( ˙ \), and diaeresis  \( ´ \) are diacritics used in phonetic spellings – as in \([pân]\) pan, \([pân]\) pain, \([kêt]\) caught and \([fâther]\) father. See Vowel Sounds and IPA;

Today several words have more common variant spellings without their original diacritic marks – for instance:

\[
\begin{align*}
\text{boutonnière or boutonniere} & \quad \text{fiancée or fiancee} \\
\text{canapé or canape} & \quad \text{garçon or garcon} \\
\text{château or chateau} & \quad \text{ingénue or ingenue} \\
\text{cliché or cliche} & \quad \text{matinée or matinee} \\
\text{cortège or cortege} & \quad \text{naïve or naive} \\
\text{coupé or coup} & \quad \text{naïveté or naivete} \\
\text{crêpe or crepe} & \quad \text{negligée or negligee} \\
\text{discothèque or discotheque} & \quad \text{noël or noel} \\
\text{entrée or entree} & \quad \text{protégé or protege} \\
\text{façade or facade} & \quad \text{suède or suede}
\end{align*}
\]

Over centuries of evolution English has come to use context, or letter order, diacritically to distinguish certain vowel and consonant sounds. For instance, in Modern English silent final \(<e>\) is used to indicate the sounds spelled by preceding letters. It marks

- **long vowels** as in mate, mete, mite, mote, mute,
- **soft \(<ç>\)’s and \(<g>\)’s** as in lice and rage, and
- **voiced \(<th>\)** as in the verbs breathe and teethe.

The letters \(<i>\) and \(<y>\) are also used diacritically to mark soft \(<ç>\) or \(<g>\) – as in lacing and lacy, and in digit and elegy. For some diacritic complications with silent final \(<e>\) see \([r]-\)Colored Vowels.

In the same vein VCC, VCV, and other **strings** are also broadly diacritic, as is a procedure like **twinning** used to maintain those strings.

**Dialects in America.** Though written American English is quite standardized, there are several dialects of the spoken language that are marked by differences in pronunciation, vocabulary, and to a lesser extent word and sentence structure. Dialects are defined by the areas from which speakers come and thus are not the same as **registers**, which are defined by the social uses of language.
American English pronunciation is still changing and evolving. A minor instance of this change is spelling pronunciation where the pronunciation of words is changed to bring it more in line with their spelling – as in often with a [t] sound. More generally there is a relentless leveling and blurring of dialect differences as American English speakers move from one geographical area to another. Blurring and leveling is also encouraged by the effects of mass communication.

Consequently dialect boundaries are fluid and hard to define, so much so that it is a real question as to how many American dialects there actually are. Some scholars speak of these basic four dialect areas:

- North, the northern third of the country from the Atlantic to the Pacific
- Coastal South, the coastal regions of Virginia, Georgia, the Carolinas, Alabama, Mississippi, Louisiana, and eastern Texas
- West, the western eleven states
- Midland, roughly everywhere else

But within these four large dialect areas more localized ones are often described, such as:
- Eastern New England
- Inland Northern
- Mid-Atlantic
- New York City
- North Central
- Southern
- Western Pennsylvania

Other dialects often described include African-American, Cajun, Chicano, and Hawaiian, and several more local dialect areas such as Appalachia, Boston, California, the Outer Banks, Maine, New Mexico, New Orleans, the Pacific Northwest, Philadelphia and Baltimore, Rhode Island, Texas, and the Upper Peninsula.

Dialects should not create many problems for teachers of spelling unless the differences in pronunciation are great enough to produce something like dialect spelling, or eye dialect, in which words are spelled so as to imitate the speller’s dialect. Intentional dialect spelling is common in literature and comic strips – thus, Li’l Abner’s Appalachian ah’ll (I’ll), yo (you), and whar’s (where’s), and Pogo’s yassuh (yes sir) and gone (going to) from his swamplands of the Coastal South. (For more on problems for students with strong dialects, see “Teaching English Spelling to Spanish Speaking Students” at “Orthographic Odds and Ends” in the Short Articles venue of this site. For more on American dialects in general see “Dialect” and “Dialect in America,” OCEL, 289-296.)
Dictionaries. I have not had enough experience with elementary and secondary school dictionaries to offer much solid advice. My one general suggestion would be to choose dictionaries that have clear and systematic treatments of etymology. The American Heritage Student Dictionary meets that criterion. And of course so does the college American Heritage Dictionary, the one abridged dictionary I have used the most and rate the highest. See also the Oxford English Dictionary and On Dictionaries and Other Helps for Teaching Vocabulary and Spelling.

Digraphs and trigraphs. Digraphs and trigraphs are strings of two or three letters that spell a single vowel or consonant sound. Some are very rare; several are doublets, doublet equivalents, or simplifications of earlier spellings. Some spell many different sounds; many deal with various pronunciations of [r]:

Vowel Digraphs:
<ae> larvae, anaerobic
<ah> hurrah, messiah
<ai> sail, said, bargain, aisle, villain
<au> taught, laugh, gauge, chauvinism, glaucoma, authority
<aw> saw
<ay> pay, says, aye
<ea> bread, bead, break
<ee> see, been, Beethoven
<ei> height, veil, neither, foreign, foreigner
<eo> leopard, people, yeoman, pigeon
<eu> eureka, neutral, neurosis, pasteurized
<ew> new, sew, few
<ey> key, prey, geyser
<ia> carriage, parliament, pneumonia
<ie> achieve, die, friend, patience
<io> fraction, ratio
<iu> Belgium
<oa> foam, broad
<oe> toe, amoeba, does
<oh> John, ohm
<oo> school, Roosevelt, flood, wood
<ou> out, poultry, thought, rough, routine, could, courteous
<ow> cow, knowledge, rainbow
<oy> toy
<ue> blue
<ui> build, suit
<uu> vacuum
<uy> buy

Vowel Trigraphs:
Consonant Digraphs:
<ar> tariff, rare, dictionary, hard, reward, coward
<bb> cabbage
<bp> subpoena
<bt> doubt
<cc> accuse
<ch> church, school, chutzpah
<ck> attack
<cq> acquaint
<cz> czar
<dd> ladder
<dg> grudge
<dj> adjective
<er> wherever, sheriff, era, sergeant, term, customer
<ff> offense
<ft> soften
<gg> struggle, exaggerate
<gh> laugh, ghost, neighbor
<gm> paradigm
<gn> sign
<ir> desirable, circle, admiral
<kn> know
<ld> should
<lf> calf
<lk> chalk
<ll> yellow
<lm> calm
<ln> Lincoln
<mb> bomb
<mm> comma
<mn> column
<ng> sing
<nn> connect
<ol> colonel
<or> scorn, word, creditor
<pb> cupboard
<ph> phone, shepherd
<pp> happy
<ps> psychology
Consonant Trigraphs:

- **<ahr>** Fahrenheit
- **<air>** pair
- **<arr>** carriage, starred, warrant
- **<aur>** dinosaur
- **<cht>** yacht
- **<ear>** wear, dear, heart, earl
- **<eer>** deer
- **<eir>** heir, weird
- **<eor>** George
- **<err>** strawberry, err
- **<eur>** entrepreneur, amateur, pasteurization
- **<ght>** flight
- **<ier>** pier, fiery, soldier
- **<irr>** irritate, squirrel
- **<oar>** roar, cupboard
- **<oor>** floor, poor
- **<orr>** worry
- **<our>** source, courage, tour
- **<rrh>** cirrhosis
- **<sth>** isthmus
- **<tch>** hatch
- **<uar>** guarantee
- **<uor>** fluorocarbon
- **<urr>** burrow
For more on vowel and consonant di- and trigraphs, see the Spellings to Sounds data table in the CommonWords database.

**Digraphs <ch, gh, ph, rh, sh, th, wh>**. The digraph <ch> was used by the Romans to transliterate the Greek chi, <χ>, spelling the sound [k], which it still does in some words like scholarly and Christmas. The <ch> was seldom used in Old English to spell [ch], which was usually spelled <ce>. The use of <ch> to spell [ch] was introduced by the Norman scribes. See Palatal sounds and their spellings. AES, pp. 356-58, 412-15.

<gh>. The digraph <gh> has a complicated history. The Norman scribes introduced it to spell the velar and palatal fricatives – that is, fricative sounds pronounced far back in the mouth, much like the sounds in the Scottish pronunciation of loch and the German pronunciation of Bach. In Old English these sounds were spelled with the rune yogh, <ʒ>, or with <g> or <h>. In time these fricative sounds fell out of spoken English, but their spellings remained.

Then the complications begin: Over time these fricatives either changed to other sounds or fell silent. As they did so, the vowels preceding them tended to lengthen or become diphthongs. As a result today the <gh>'s that descend from Old English usually follow a long <i> or another long vowel or diphthong spelled <ai> or <ei>, or <ou>:

- flight, high, knight, light, night, right, sigh, sight, thigh;
- straight; eight, freight, neighbor, sleigh, weigh, height;
- dough, though, thorough; through; bough, plough.

In a few cases the <gh> follows the low back vowel [o] spelled <au> or <ou>:

- caught, daughter, haughty, naughty, slaughter, taught;
- bought, brought, fought, ought, sought, thought.

And in a few cases <gh> spells [f] after <au> spelling [ä] or <ou> spelling [ü]: draught, laugh, laughter; enough, rough, slough "discard", tough.

Other than the native burgh, none of the <gh> spellings of [g] descend from the Old English fricatives. Ghost is probably a Flemish-influenced respelling of earlier gost by the first English printer William Caxton, and aghast and ghastly are probably respellings via analogy with ghost. The remainder are foreign adoptions — afghan, dinghy, ghetto, ghoul, sorghum, spaghetti, yoghurt, etc. (AES, pp. 209-10, 237-38, 379-80, and passim).

<ph>. The Romans used <ph> to transliterate the Greek phi, <φ>, which spelled a sound close to the Roman [f]. In time the two sounds became confused. There are a few reminders of that earlier confusion in
variants like *sulfur, sulphur* and in historically related groups of words like *fantastic, fantasy; phantasm, phantom*. Most English words with <ph> are from Greek: *alphabet, apostrophe, atmosphere, biography, chlorophyll, elephant, emphasize, geography, hyphen, metaphor, orphan, paragraph, philosophy, phone, physical, prophet, sphere, triumph*. AES, pp. 378-79.

**<rh>**. The Romans used <rh> to transliterate the Greek rho, <p>, which spelled a sound close to the Latin [r]. In time the Greek and Latin sounds converged to [r]. The digraph <rh> is fairly common in the scientific-technical register: *rhapsody, rheostat, rhetoric, rheumatic, rhinoceros, rhomboid, rhododendron, rhubarb, rhythm*. Double rho was transliterated as <rrh>, as in *arrhythmia, catarrh, diarrhea, gonorrhea, hemorrhage, myrrh*. AES, pp. 448-49.

**<sh>**. The Norman scribes introduced the <sh> spelling of [sh] to replace the Old English <sc>. Since they were not familiar with the [sh] sound, they tried several ways of spelling it, and <sh> was finally settled upon in the 15th century. See *Palatal sounds and their spellings*. AES, pp. 409-09.

**<th>**. In Old English the runes edh, <ð>, and thorn, <þ>, spelled both voiceless [th] as in *thin* and voiced [th] as in *then*. In late Old English or early Middle English edh dropped out, and thorn was used to spell both [th] and [ð]. Then in the late 14th and early 15th centuries the Norman scribes gradually replaced thorn with <th>, again to spell both sounds. AES, pp. 384-86.

**<wh>**. In Old English <hw>, spelling [hw], appeared in several words: *hwæl* “whale,” *hwæt* “what,” *hwēol* “wheel” – and about 200 others with initial <hw> in J. Hall’s *A Concise Anglo-Saxon Dictionary* (Cambridge: 1975). The Norman scribes inverted the spelling, perhaps via analogy with <ch>, <sh>, <th>. Today we have modern words with <wh> with their earlier pronunciations with [hw], sometimes reduced tho [ˈw], and often reduced further to [w]. In a few words <wh> has come to spell [h]: *who, whole, wholly, whom, whose*, for instance. AES, pp. 387-89, 458-59.

**Diphthongs.** The word *diphthong* combines two Greek elements: *di-* “two” and *phthong* “sound.” A diphthong runs together two vowel sounds. Our analysis recognizes only two modern English diphthongs: [oi] spelled <oy> and <oi> as in *coy* and *coil*, and [ou] spelled <ow> and <ou> as in *how* and *house* (in IPA /ɔɪ/ and /aʊ/ respectively). The [oi] diphthong is spelled <oi> about 2:1 over <oy>, which almost always occurs at the end of free stems. The diphthong [ou] is spelled <ou> about 70% of the time, <ow> about 30%. It is rarely spelled <au>, as in *glaucoma*, one pronunciation of *trauma* – and *Schnauzer*.

Although our analysis recognizes only two diphthongs, different speakers – especially those for whom English is a second language –
often diphthongize our [ɪ] to /aɪ/ and to a lesser extent our [æ] to /eɪ/ and [ə] to /əʊ/ or /ou/.

**Doublets and doublet equivalents** are consonant digraphs (and the trigraph <tch>) that spell a single consonant sound, usually after a short vowel, and usually in word-medial position due to the **twinning** of a final consonant when adding a suffix as in **twinning twin+n+ing)** or the **assimilation** of final consonants in prefixes as in **announce (ad+n+nounce** and **acquire (ad+c+quire**:

- <bb> = [b] as in robber
- <cc> = [k] as in accurate
- <ck> = [k] as in rock
- <cq> = [k] as in acquire
- <dd> = [d] as in reddest or odd
- <dg> = [j] as in bridge
- <dj> = [j] as in adjourn
- <ff> = [f] as in offer or off
- <gg> = [g] as in bigger or egg
- <ll> = [l] as in allow and tell
- <mm> = [m] as in hammer
- <nn> = [n] as in dinner or inn
- <pp> = [p] as in happy
- <rr> = [r] as in carry
- <ss> = [s] as in missing or kiss
- <tt> = [t] as in cotton
- <zz> = [z] as in dizzy or fuzz

**Doublets Misspelled.** That old standby **Mississippi** has three doublets, which is pretty unusual. But several words contain two doublets, and it can be hard to remember them both. **Explication** can usually show where those doublets come from and why they’re there – which is nearly always due to either **assimilation**, **twinning**, **simple addition**, or the **VCC** pattern:

- **accommodate** (ad+c1+(com+mod4+ate)1
- **aggression** (ad+g1+gress+ion)1
- **embarrass** (en1+m2+barr3+ass) The suffix -ass) marks French verbs.
- **dissolve** (dis+solve
- **goddess** god1+d2+ess)1
- **Halloween** hall2+ow)+een The base een contracts evening.
- **misspell** (mis+spell1
- **occurred** (ob+c1+cur1+r2+ed)1
- **possession** pos1+sess+ion)1 “The act of sitting as master”
- **successful** (sub+c1+cess1+ful)

Some words contain a doublet and a singleton, and it can be hard to remember which is which:

- **Caribbean** Carib+b+ean)
- **disappear** (dis+(ad+p1+pear2
- **disappoint** (dis+(ad+p1+point
- **harass** har1+ass)
Sometimes it can be hard to remember that you need a doublet:

- **finally** (fin4+al)1+ly)1
- **interrupt** (inter+rupt
- **preferred** (pre+fer1+r2+ed)1
- **really** re1+al)1+ly)1
- **referred** (re+fer1+r2+ed)1

Nearly all instances of the doublet <ss> occur in just four bound bases – gress (in 98 words in Lexis), miss3 (128 words), sess (76), and cess1 (165) and in common free bases like class1 (83), cross (93), pass1 (95), press (240).

**Early Modern English.** Early Modern English is English as it was spoken and written from roughly 1450 to 1700. It is the English of Caxton, Shakespeare, Mulcaster, Jonson, Marlowe, Spenser, Donne, Milton, and Dryden. It was the period during which English spelling became increasingly standardized as part of a [Standard Written English](#).

**Elements.** Elements are the written counterpart of the morphemes of the spoken language. They are the smallest parts of written words that have the following two features:

(i) In words in which they occur, they contribute the same or closely related content.

(ii) They are spelled consistently from word to word – that is, spelled either the same or with variants that can be rationalized, as with the <aesth> and <esth> variant spellings in aesthetics and esthetics.

There are four kinds of elements: free bases like plant and car; bound bases like fect and fer as in perfect (per1+fect and defer (de+fer1; and prefixes and suffixes, like (re- and -ed)1 in replanted (re+plant+ed)1.

Elements often come in sets of two or more variants – for instance, +arter and its expansions +arteri and +arterio “windpipe, artery”, as in artery, arterial, arteriosclerosis. Elements also occur in what I call co-sets: pairs, or rarely triplets, of elements that are related, but not necessarily equivalent, semiotically and that operate as a team in different parts of speech – for instance, the co-set {+ceed, +cess} as in the verb succeed and its noun of result success. Another example is the triplet set {+tain1, tent2, ten2}, as in
• abstain (abs+tain1, abstention (abs+tent2+ion1);
• contain (com+n1+tain1, contents (com+n1+tent2+s)8;
• detains (de+tain1+s)2, detention (de+tent2+ion)1;
• maintain main2+tain1, maintenance main2+ten2+ance);
• retain (re+tain1, retentive (re+tent2+ive); and
• sustain (sub6+s1+tain1, sustenance (sub6+s1+ten2+ance), sustention (sub6+s1+tent2+ion)1.  AES, pp. 39-47

**Eminent, imminent.** Both of these words contain the base min2 “to project, jut, threaten.” The differences are in the prefixes: (ex+min2+ent) and (in+m+min2+ent). The (ex- is an assimilated form of (ex- “out, outside”: Something that is eminent juts out or shows – above others. The prefix (in2- means “in, within”: Something that is imminent is hidden but can soon project itself or jut out.

The different treatment of (ex- and (in2- occurs in several pairs of words starting <em> and <imm>: emigrate, immigrate; emigrant, immigrant; emerge, immerge; emersion, immersion.

**Envelope, envelop.** Envelope is the common noun referring to the paper cover you put mail into; envelop is the less common verb meaning “to wrap or enclose.” They are the same word with slightly different spellings and pronunciations to mark the difference in part of speech. Notice that the spelling of the final syllable of envelope is an example of the VCV pattern with the long <o>.

**-er)01, -or)2.** These two agent suffixes form nouns that refer to persons or things that perform actions: A teacher teaches and a director directs. There is no reliable way to tell when to choose one and when the other. One hint is that -er)01 is five times more common in Lexis than is -or)2. In CommonWords it is only about twice as common. When in doubt, a good dictionary is your best bet.

**Essay, assay.** Essay explicates to (ex+s1+say2, and assay explicates to (as2+say2. Historically they are the same word, whose spellings and pronunciations have shifted back and forth over the centuries, and have settled down so that essay as both noun and verb has the sense of a try or trial, including that try that leads to piece of short expository prose. Assay is pretty much restricted to the evaluation of metals. The prefix (as2-alters (ex- “out, away,” and the base say2 originally meant “drive, draw, move” – but in Latin meant “to weigh.”

**Etymology.** Etymology is the study of the origins of words and the routes they took getting into English. The etymological information in a good desk dictionary like the *American Heritage Dictionary* or the *American Heritage Student Dictionary* can help composition students think of things to say about their topics as they find unifying connections among what at first may have seemed to be unrelated words and ideas. It can help them keep their
minds on their topics, see useful analogies, and create pertinent metaphors. And it can help them develop their reasoning skills, as they sort, discriminate, and link the ideas and forms they find in their words’ histories.

Working with etymology can also help spelling students see some of the connections that hold our spelling system together. These connections can help increase a sense of ruliness and order in our spelling – and thus ease and strengthen learning. For more on etymology and its uses in the classroom go to “Some Uses of the Sources Field” in Introduction to CommonWords, also Lineages from Indo-European Roots and Why Study Indo-European Roots.

**Evolution.** Back in the Middle Ages English spelling was a tangle of variant spellings, driven primarily by wide differences in dialects and pronunciation. For instance, the OED lists about 50 different early spellings of the relatively simple word *harbor*. But as the language evolved from Old to Middle to Early Modern English, that variation gave way to a high degree of standardization – enough so that by 1600 nearly every English word had one and only one standard spelling. See Standardization in Early English Orthography. The following is an example of word endings evolving to be more similar to one another in order to highlight their parallel functions: Old English *mōdor*, *brōthor*, *fæder*, and Old Norse *systir* evolve to *mother*, *brother*, *father*, *sister* – all with the common \[cɛr\] ending spelled <er>.

That evolution to regularity continues as foreign adoptions are adapted to fit the English spelling system – see, for instance, foreign noun plurals – and as nonregular native spellings are rendered more regular. For instance, verbs that traditionally indicate the past tense with -t)1 often acquire a variant with the regular -ed)1: *burn*, *burnt*, *burned*; *dwell*, *dwelt*, *dwelled*; *kneel*, *knelt*, *kneeled*; *spell*, *spelt*, *spelled*. British spelling tends to be more conservative than American towards such changes, consistent with the differences in attitude between Samuel Johnson and Noah Webster. See British and American spelling and Metaphor.

**Excel, excellent.** From the modern point of view the major question here is “Why the <ll> in *excellent*? Though the <cel> is stressed in *excel*, in *excellent* it is not. Thus, *excellent* would appear to be a holdout to the Twinning Rule, which would require stressed <cel> in both words. From the historical point of view a better question would be why isn’t *excel* spelled <excel>? Both *excel* and *excellent* come from Latin by way of French, and both languages had <cell>. In fact, *excel* was regularly spelled *excell* up through the 17th century. The OED’s earliest citation of *excel* is in the 18th century.

Actually, the contraction of <ll> to <l> has been quite common in English, especially in stressed final syllables. For instance, the variants *fulfill* and
fulfil are from full plus fill. In the more common spelling, fulfill, the first <ll> is contracted, and in the variant fulfil both are. Similarly the more or less archaic withal “besides, likewise, therewith” and its compound wherewithal are from with plus all and were often spelled with <ll>. See All and Its Compounds. Similar contractions have occurred in, among others, compel, Brazil, parallel, cartel, rebel and until (vs. till) – all of which had earlier <ll>. See AES, pp. 440-41.

**Except, accept.** These two words are close enough in structure and pronunciation to create problems for spellers. They both contain the base cept “grasp, take, seize.” The prefixes tell the story: Accept contains the prefix (ac1-, an assimilated form of (ad- “to, toward.” And when you accept something, you do take it toward you. Except contains the prefix (ex- “out, away.” And when you except something, you take it away or leave it out.

There are other pairs of words that have some form of the prefixes (ad- and (ex-: with the same base:

- **access** (ad+c+cess1 vs. **excess** (ex+cess1 and **accede** (ad+c+cede vs. **exceed** (ex+ceed: The bases cess1, cede, and cede’s rare variant cease mean “go, yield, come.” When you access something, you basically go to it or come to it, but an excess is a going away, in the sense of going beyond some limit. To accede means to come to, while to exceed is basically to go beyond. Choosing between the <ceed> and <cede> spellings can be helped by this little mnemonic: “If you proceed and do not exceed, you will succeed,” which lists the three instances of <ceed>.

- **accuse** (ad+c+cuse vs. **excuse** (ex+cuse: The base cuse means “reason, purpose.” To accuse someone of something is to assign a certain reason or purpose to him, while to excuse someone of someone is to take that reason or purpose away from him.

- **affluent** (ad+f+flu1+ent) vs. **effluent** (ex+f+flu1+ent): The base flu1 means “flow, swell.” The adjective and noun affluent have the sense of flowing toward or swelling, while effluent means something that is flowing away.

**Explication.** Explication is an analysis meant to unfold information in written words, information that is not just useful to describing and understanding written English, but also to teaching and learning it. Explication attempts to do so by identifying the basic units in written words – that is, the elements and particles. It also shows the historical processes – such as assimilation and palatalization – and the procedures that are followed when the elements concatenate – such as simple addition, final <e> deletion, and twinning. To find information about the elements and particles in an explication, including those with index numbers, you can check the CommonWords and Lexis databases on this website. For more on explication see “The Explication of Written Words” (AES, pp. 32-66) and On Explication in the Short Articles venue of this website.
Finally, finely. Most commonly these two adverbs are homophones, pronounced [fîn'îlë], though sometimes you will hear finally pronounced with three syllables, [fîn'î-îlë]. Finally explicates to fin4+al1+ly1; finely to fine1+ly1. Both fin4 and fine1 come from a Latin word that meant “limit, boundary.” That fin4 is the same as the fin you see at the end of French movies, meaning “end.” Early on fine meant “of the highest quality” but it has come most commonly to mean “consisting of small particles.”

Folk etymology. See Metaphor.

Fossils. In English spelling one type of fossil is a letter that earlier served a function but no longer serves that function and has not acquired a new one. Most fossils are silent final <e>’s that are in their words only for etymological reasons. Most fossils are in words from French; fewer are in native words. In French words final <e> marked feminine gender, and since English does not inflect words to distinguish gender, that function was lost but even in words in which the now silent final <e> has no new function it continues in the spelling: avalanche, brassiere, cigarette, clientele, finesse, impasse, medicine, troupe. Some native English words contain fossil final <e>’s: come, done, forbade, gone, none, some. In some suffixes there are fossil silent final <e>’s when the suffix is unstressed: In the verb approximate the suffix -ate) is stressed with long [ä], but in the adjective approximate -ate) is unstressed with short [I]. Also with -ate) are words like certificate and immediate. Similar cases involve the suffixes -ite2, -ure), and those spelled <ine>: composite, definite, favorite, infinite; capture, culture, failure, pleasure; discipline, feminine, masculine.

It would be possible treat as fossils certain letters in simplifications. For instance, in the simplification <mb> as in limb, one could treat the <b> as a fossil, a letter that has lost its earlier function of spelling [b]. I have chosen not to follow that line of analysis, which seems to me to get more complicated. AES, pp. 148-54.

French influence. In 1066 England was conquered by the Normans, Vikings who had settled in northern France roughly a century earlier. Over time the Normans (or “North Men”) had adopted their version of French, and after the Norman Conquest, Norman-French became the language of power and education in England. Over the centuries, due largely to its interaction with English, the language of the common people, and with Latin, the language of the church, Norman-French evolved into the dialect known as Anglo-Norman. Although adoptions and adaptations from French go on today, a huge number of them came through Norman-French during the Middle Ages, many of them replacing native English words. Some scholars estimate that about 45% of all English words are from French. See Caxton and Norman scribes.

French Lemon (Stress Frontshift) Rule. Words In French carried stress
on the final syllable unless that syllable was a weak <e>, in which case stress fell on the next to final syllable. As French adoptions were integrated into English, the stress shifted forward, especially in nouns and in line with the native English pattern of stressing the first syllable of the base. But before the stress frontshifting occurred, French adoptions would have been pronounced with the first vowel weakly stressed and either reduced or short. Then, when the stress moved forward in the word, these vowels, now stressed, would be unlikely to lengthen. As a result there are in English hundreds of French adoptions that have short stressed vowels in patterns that would lead us to expect long vowels. This stress frontshifting preempts the more general VCV vs. VCC pattern, which stipulates that a stressed vowel letter followed by a single consonant which is in turn followed by another vowel letter will be orthographically long, as in a words like *demon* and *dame*.

However, in a French adaption like *damage*, although the stress is on the first vowel, which heads a VCV string, that head vowel is short, [ā]. And the spelling is with single <m> rather than <mm>, not *dammage*, with the VCC pattern normal for medial short stressed vowels. This preemption of the VCV vs. VCC pattern is perfectly regular in French adoptions. Some examples of the non-French and French contrast: *demon* vs. *lemon*; *yodel, model; molar, scholar; driver, river; specious, precious; navel, gravel*. For more on the Stress Frontshift Rule see AES, pp. 127-30 and *Standardization in Early English Orthography* in the Short Articles venue of this website.

**Frequentatives.** A frequentative is a word that indicates repeated action. The suffixes -el)2, -l)1,-le)1, and -er)03 mark frequentatives – as in the verbs *drivel, snivel; drawl, snarl; babble, crumble, jiggle; bicker, clatter, flutter*. Frequentative verbs are often converted to frequentative nouns – *drivel, drawl, snarl, clatter*. A few frequentative verbs are formed by reduplication of a monosyllable – as in *coo-cooing or bob-bob-bobbing*. And a few frequentative nouns can be formed by combining two different vowel sounds in the same word – as in *teeter-totter, pitter-patter, chitchat*, etc.

**Fricative sounds.** Fricatives are consonant sounds articulated with enough closure of the vocal tract to produce friction – [v] as in *verve*, [f] as in *fluff*, [tθ] as in *thin*, [θ] as in *then*, [z] as in *zebras*, [s] as in *sass*, [zh] as in *azure*, [sh] as in *shush*, [h] as in *hurt*.

**Futhorc.** The futhorc was the runic alphabet used by the Anglo-Saxons. See *Runes*.

**Genitives.** See *Possessives*.

**Germanic languages.** Germanic is one of the 15 or so branches of
languages in the **Indo-European language** super-family. The Germanic branch is traditionally divided into three: North Germanic, which includes Icelandic, Faroese (spoken in the Faroe Islands of the north Atlantic), Norwegian, Swedish, and Danish; West Germanic, which includes English, Frisian (spoken in the Frisian Islands of the North Sea, the language closest to English), Flemish (the Dutch language as spoken in Belgium), Dutch, Afrikaans (spoken in South Africa), German, Yiddish; and East Germanic, which includes only the now extinct Gothic. There is no Southern Germanic branch. See [Norse and Anglo-Norse](#) and “The Family of Languages” in the [Lineages from Indo-European Roots](#) venue in this website.

**Gerunds.** A present participle used as a noun.

**Glides.** A glide is the transitional sound created when moving from one sound to another. One such glide is usually spelled <i>, rarely <y> or <e> – as in azalea, canyon, civilian, million, onion, savior, spaniel.

**Glottal sounds.** Glottals are *consonant sounds* produced well back in the mouth with closure at the glottis at the top of the vocal cords. The only modern English glottal sound is [h].

**Great Vowel Shift.** The Great Vowel Shift was a huge shift that occurred in the pronunciation of English vowel sounds, especially long vowels. It started around 1350 and was for the most part finished by 1600, though changes do continue. It is a complicated affair and not entirely understood, but we can make just two general points:

First, in *Middle English* long vowels had Continental values – that is, the sounds we hear today in Italian and Spanish. For instance, the long <i> in Middle English was pronounced [ē] as it does in, say, *burrito*; the long <a> was pronounced [ä] as it is in *sonata*. Today due to the shift, long <i> in English is pronounced [í] as in *bite*, and long <a> is pronounced [ä] as in *bait*. These and similar shifts account for much of the very different sound of English as compared with other European languages.

Second, since the spoken language changes much faster than the written, and since English spelling was becoming standardized in late Middle and *Early Modern* English (see [Standard Written English](#)), modern spelling in some respects does a better job of spelling earlier English than the English of today.

**Holdouts.** For two reasons I prefer the word *holdout* to the overused, and usually pejorative, *exception*. First, *exception* has acquired a bad reputation in English spelling, and is often used in the argument that our spelling system is ramshackle and essentially hopeless, which it is not. Folks tend to be too quick to label some spelling as an exception. Second, *holdout* does a better job of suggesting that in time we may learn enough about English spelling that this spelling can be treated as regular. And
Holdout better fits the step-by-step strategy that I advocate for dealing with spellings and whose goal is to make any spelling as rational and systematic as possible:

Step one: Try to include it within a pattern, procedure, or process that is universal and ruly in the system. For instance, to explain the <tt> spelling of [t] in, say, lettuce, treat it as an instance of the tactical contrast VCV vs. VCCV.

Step two: Failing step one, include the spelling in a more localized subpattern or subprocess, which can preempt a more general rule. For instance, to explain the <ck> spelling of [k] in picnicking, treat it as an instance of the doublet equivalent <ck> used when adding -ing1 to picnic, since normal twinning would lead to *picniccer, pronounced with [ks], not [k].

Step three: Find a historical reason for the spelling – for instance, treat the final <b> in the word bomb as a simplification. And notice, too, that sometimes you can still hear the <b> in bomb – for instance, in bombadier or bombard.

Step four: Failing steps one, two, and three, list the spelling, describing its irregularity and relating it to larger, more rational patterns and rules within the system – for instance, the <or> spelling in forty, as contrasted with four, fourteen, fourth. Historical note: Before the 15th century the spelling was commonly <our>. Forty appeared in the 15th century and dominated thereafter. Why four went one way and forty another is not clear, but it may have something to do with the fact that each spelling is a simplification of the Old English fêower “four” and fêowertig “forty.”

Homonyms, homophones, homographs. Basically the words homonym “same name” and homophone “same sound” refer to two or more words that have the same sound, sometimes the same spelling, but different senses – for instance, bear1 “to carry” and bear2 “large mammal,” or bear1, bear2, and bare. Homograph “same letters” refers to two or more words that have the same spelling but different sounds and senses: wind1 “moving air” pronounced [wɪnd] and wind2 “to wrap” pronounced [wɪnd]. Homonyms and homophones pose problems for spellers – pairs like alter, altar and council, counsel can be tough. And CommonWords lists 747 words that have homophones. See also Synonyms, antonyms, pseudonyms, heteronyms, acronyms

Hyphen. A hundred or so years ago I took a course in grammar from Porter Perrin, a delightful, pipe-smoking, gentle man. One day a distraught student asked him about the use of hyphens. Dr. Perrin turned and stared out the window, took a few puffs on his pipe, and said, wistfully, “Anyone
who takes the hyphen seriously shall surely go mad,” a paraphrase from the Oxford University style book. A look at the three-and-a-half page discussion of the hyphen in Dr. Perrin’s *Writer’s Guide and Index to English* (Scott, Foresman: 1950) confirms that judgement. But the following may help some. Use a hyphen:

- to mark the division of words at the end of a line
- after prefixes that end with the same vowel letter with which the stem begins: *re-elect*
- to avoid confusion with other words: *re-cover* “cover again” vs. *recover* “get back”
- in written-out numbers from *twenty-one* to *ninety-nine*
- in names of *in-law* family relationships: *mother-in-law, sister-in-law*
- in some other **compound words**: *manic-depressive, so-called, well-known*

In general modern English uses fewer hyphens than it did in the past. So the best advice here remains “Check a current dictionary.”

**<i> before <E>**. Perhaps the best known mnemonic device for English spelling is “It’s *<i>* before *<e>* except after *<c>* or when sounded [ä] as in *neighbor* or *weigh*.” And it’s a useful jingle: When spelling [ë], it is regularly *<i>* before *<e>* except after *<c>*. Thus it is *<ie>* in *achieve, belief, grieve, hygiene, piece, shriek*, but *<ei>* in *ceiling, deceit, perceive, receive*. The only known holdouts with [ë] spelled *<ei>* after something other than *<c>* are *protein, seize, weir, weird* – and *either, neither, leisure*, the last three of which have variant pronunciations without [ë].

The jingle also says that it is *<ei>* rather than *<ie>* when you are spelling [æ], a generalization with no known holdouts: *eight, neighbor, reign, vein, weigh*.

Expanding on the jingle somewhat, when you are spelling [i], it is *<ie>* in final position (*die, hie, lie, pie, tie, vie*), but *<ei>* in initial and medial position (*eiderdown, height, kaleidoscope, poltergeist, seismic, sleight, stein*). The only known holdouts are *fiery* and *hierarchy*, with *<ie>* in medial position.

So far as spelling short vowels is concerned, the jingle holds quite well, the only known holdouts being *counterfeit, foreign, forfeit, heifer, sovereign, surfeit*—all of which have *<ei>* after something other than *<c>* (and three of which contain the same bound base, +feit). The number of holdouts is quite modest in view of the hundreds of words that are covered by the jingle, especially if it is expanded to cover long *<i>*.

**Illusion, allusion.** The base of these two near homophones is *lus1*, which originally meant “to play, jest.” It is related to the *lud(e)* in *ludicrous, interlude, and prelude*. *Allusion* came to us by way of a Latin word which kept that original sense, but *illusion* comes from a Latin word with a more pejorative sense, “to mock.” You have to pay special attention to the first vowel in each word: *<i>* pronounced [i] in *illusion*; *<a>* pronounced [a] in *allusion.*
**Imitative.** In Lexis and CommonWords *imitative* is used to refer to what is sometimes called onomatopoeic or echoic words. An imitative word sounds like the thing it refers to: *beep, click, hum*, etc. There are 83 imitative words listed in CommonWords, in Lexis 454. With some imitative words it is hard to say what exactly is being imitated: *boudoir, gallimaufry, obstreperous*.

Imitative language is a mainstay in poetry, even though the individual words may themselves not be particularly imitative – as in these lines from Mathew Arnold’s “Dover Beach” describing the sound of the waves washing up on the shore of southern England:

Listen! you hear the grating roar  
Of pebbles which the waves draw back, and fling,  
At their return, up the high strand,  
Begin, and cease, and then again begin,  
With tremulous cadence slow, and bring  
The eternal note of sadness in.

The notion that all language originated in the imitation of natural sounds – the so-called bow-wow theory – is now mostly discredited, but surely natural sounds are the source of several of our words.

**Indo-European languages.** English is one of several hundred languages in the Indo-European super-family, which descends from Proto-Indo-European and includes languages in the Slavic, **Germanic**, Celtic, Italic, Hellenic, Anatolian, and Indic families, and several others not represented in CommonWords. Today more than 40% of the people of the world speak an Indo-European language as their native tongue.

In the chart below languages appear unboxed, the Romance sub-family in a very light gray box, families in darker grey, the Indo-European super-family in darkest gray.)

Proto-Indo-European, the mother tongue of the Indo-European super-family, is thought to have been spoken around 5000 B.C. in the area north of the Black and Caspian Seas. Over the millennia the Indo-Europeans spread east to India and central Asia, west to modern Greece, Italy, France, Spain, south to Iran, Pakistan and Afghanistan, and north to Germany, Britain, and Scandinavia. The spread continues today, especially as English becomes an increasingly international language.
Indo-European Roots. Indo-European roots are meaningful units, or morphemes, that have been reconstructed from the Proto-Indo-European language (PIE), which we believe was spoken north of the Black and Caspian seas more than 5,000 years ago. PIE is the source of the Indo-European family of hundreds of languages, including English and the more than 400 that are still spoken today.

PIE was not a written language, so all we know of it has been reconstructed over more than two hundred years of study of sound patterns and senses among its descendant languages. The reconstruction of these roots – so far, well over a thousand – and identifying their descendant words is one of the great linguistic detective stories. Some examples:

<table>
<thead>
<tr>
<th>Root</th>
<th>Sense</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>*beu</td>
<td>“To swell”</td>
<td>boast, bosom, puff, pocket</td>
</tr>
<tr>
<td>*bheu(e)</td>
<td>“To be, exist, grow”</td>
<td>be, bond, husband, physics</td>
</tr>
<tr>
<td>*dwo</td>
<td>“Two”</td>
<td>balance, between, twelve,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>two</td>
</tr>
<tr>
<td>*gene</td>
<td>“To give birth, beget”</td>
<td>general, genetics, kin, kind,</td>
</tr>
<tr>
<td>*gnō</td>
<td>“To know”</td>
<td>naive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>could, know, notion,</td>
</tr>
</tbody>
</table>
An accessible and highly thorough treatment of Indo-European roots as they are reflected in the modern English lexicon is that of Calvert Watkins and his group. Their work can be found as the appendices of the 1st, 3rd, 4th and 5th editions of the American Heritage Dictionary and in the three editions of his paperback The American Heritage Dictionary of Indo-European Roots (1985, 2000, 2011), both of which have changed from one edition to the next to reflect new editorial work.

Watkins’ work like that of all students of Indo-European roots relies on Julius Pokorny’s magisterial study The Indogermanisches etymologisches Wörterbuch (Indo-European Etymological Dictionary) (1957, 2005), several different versions of which are available in print and on the Internet. There are also several on-line versions of Indo-European roots.

For an annotated map of the Indo-European family, together with examples of some of the roots, the thousands of modern English words that descend from them, and some reasons for having students work with roots, see Lineages from Indo-European Roots elsewhere on this site.

**Inflection, inflected forms.** Inflected forms are words formed by adding inflectional suffixes to nouns, verbs, adjectives, or adverbs.

**Innocence, innocents.** Innocence (in1+noc+ence) refers to the quality of being innocent, its -ence) suffix marking it as an abstract noun. The -ent) suffix in innocents (in1+noc+ent)+s3) forms adjectives and non-abstract nouns. The base noc “death, harm, injure” is negated by the prefix (in1. So an innocent isn’t likely to hurt you. The issue here is to distinguish between the abstract noun on one hand and the homophonic non-abstract plural noun on the other. There are actually a number of -ence) and -ent) pairs in English – for instance, adolescence (ad+ol1+esc)+ence), adolescents (ad+ol1+esc)+ent)+s3; fluorescence flor1+escē)+ence), fluorescent flor1+escē)+ent)+s3; magnificence magni+fic)+ence), magnificents magni+fic)+ent)+s3. But they shouldn’t create the problems that innocence and innocents can, because beyond the three listed above, very few of the pairs have a plural noun ending in <ents>. In fact you have to have a pretty big dictionary to find magnificents.
Insulators. There are two letters that routinely serve to cover, or insulate, other letters. The most common is silent final <e>, which insulates <s, z, u, v> at the end of words – as in lapse, bronze, plague, serve. See silent final <e> functions and deletion for more details. In some <gu> spellings the <u> insulates a following <e, i, y> to avoid the look of a soft [j] sound: guest, plaguing, plaguy. See soft and hard <c> and <g>.

Intensifiers. An intensifier is a word or word part that increases the intensity of the word it modifies. In English the most common intensifier is the word very. Here are a few of the several dozen English intensifiers, with example phrases: awful “awful rich,” bloody “bloody amazing,” dead “dead wrong,” real “real tall,” really “really cold,” so “so stupid,” super “super hot,” too “too tense,” wicked “wicked smart” – plus many -ly1 adverbs like extremely.

Lexis lists 35 prefixes that are tagged as rarely or sometimes or often intensifiers. But in nearly every case the intensification is hard to discern in modern words – perhaps -per1 in perfect and perceive and (be- in belief and believe would be instances. Also two suffixes are tagged “intensifier”: the quite productive -ard1 as in blizzard, drunkard, and the less productive -ee4 as in whoopee, sirree. And we use some prepositions after verbs to intensify the action: there’s a difference between merely burning something and burning it up or down.

Interdental sounds. Interdentals are consonant sounds pronounced with the tongue between the upper and lower teeth – [th, th] as in then and thin.

International Phonetic Alphabet (IPA). The IPA is a phonetic alphabet designed for transcribing any of the world’s spoken languages. It expands the Latin alphabet, with more than one hundred consonant and vowel symbols and seventy or so other marks. In Daniels and Bright’s presentation of the 1989 version of the IPA in their The World’s Writing Systems (Oxford UP, 1996) there are 99 letters (74 consonants and 25 vowels), 31 diacritics, 11 suprasegmentals (which indicate qualities that extend beyond one segment, such as intonation patterns), 12 tone and accent marks, and a group of 18 other symbols for non-English speech sounds. The IPA is used to make both phonemic (or broad) and phonetic (or close, detailed) transcriptions. Phonemic transcriptions are enclosed in diagonals; phonetic transcriptions in square brackets. (NB: On this website the transcription of sounds is essentially phonemic, or broad, though – perversely enough – the sounds are usually enclosed in square brackets.) See consonant sounds, vowel sounds, and diphthongs, where IPA letters are shown in diagonals.

-ize), -ise)1. These two suffixes usually form verbs – and a very few nouns. In America we prefer -ize), while in Britain they prefer -ise)1, though they use both. (See British and American spelling.) About the only common American words that always take -ise)1 are advertise, chastise,
exercise, exorcise, franchise, and the noun merchandise; the verb merchandise can take either. The Lexis database lists 69 words with -ise)1, 2,875 with -ize). When in doubt, -ize) is clearly the better bet.

Jespersen, Otto. Otto Jespersen (1860-1943) was a Danish linguist who studied and wrote extensively about English grammar. His many works include Growth and Structure of the English Language (1905), Language: Its Nature, Development and Origin (1922), The Philosophy of Grammar (1924), and Modern English Grammar on Historical Principles (7 vols., 1909–1949), the first volume of which, Sounds and Spellings, is particularly useful to students of English spelling.

Johnson, Samuel. Samuel Johnson (1709 – 1784) was an English writer who made lasting contributions to English literature as a poet, essayist, moralist, literary critic, biographer, editor and lexicographer. He is sometimes said to be the most distinguished English man of letters. His 1755 A Dictionary of the English Language was the first true English dictionary and set the standard for British English spelling. (See British and American spelling.) At times his sense of humor showed through, as when he defined a network as “any thing reticulated and decussated, at equal distances, with interstices between the intersections,” or when he defined oats as “a grain, which in England is usually given to horses, but in Scotland supports the people.”

Labiodental sounds. Labiodentals are consonant sounds pronounced with the lower lip close to or touching the upper teeth – [v, f], as in verve and fife.

Language Families. The figures are uncertain and fluid, but there are 7,000 to 7,500 known languages spoken in the world. Of these only a fraction have fully developed and widely used writing systems. Thousands of languages are spoken and heard; far fewer are written and read.

The uncertainly and fluidity multiplies when linguists try to group these thousands of languages into the 150 or so proposed language families – that is, groups of languages thought to descend from a common mother tongue. These proposed families contain widely different numbers of both languages and native speakers and they cover geographical areas of widely different sizes, as shown in the small sample below:

<table>
<thead>
<tr>
<th>Family</th>
<th>Geographical Area</th>
<th>Languages</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afroasiatic</td>
<td>North Africa and western Asia</td>
<td>360</td>
<td>500 million</td>
</tr>
<tr>
<td>Austronesian</td>
<td>Indonesia, Madagascar, Sumatra, the Philippines, and several Pacific islands</td>
<td>1,250</td>
<td>400 million</td>
</tr>
<tr>
<td>Family</td>
<td>Geographical Area</td>
<td>Languages</td>
<td>Speakers</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>Indo-European</td>
<td>Western Asia, India, Iran, Europe, the UK, United States – still spreading</td>
<td>444</td>
<td>3.2 billion</td>
</tr>
<tr>
<td>Niger–Congo</td>
<td>Sub-Saharan Africa</td>
<td>1,500</td>
<td>430 million</td>
</tr>
<tr>
<td>Sino-Tibetan</td>
<td>China and Tibet</td>
<td>457</td>
<td>1.3 billion</td>
</tr>
<tr>
<td>Trans-New Guinea</td>
<td>New Guinea and surrounding islands</td>
<td>480</td>
<td>3.5 million</td>
</tr>
</tbody>
</table>

Some scholars have proposed an even larger super family, **Eurasiatic**, which would include Indo-European and several other languages and families, including Finnish, Korean-Japanese and Eskimo-Aleut, with strong ties to American Indian. And there are proposed super families that are even larger – with boundaries even more uncertain and fluid.

For reasons more religious than linguistic people in the past believed that English (and other modern languages) had descended from Hebrew. Now we know that English is not only not a descendant of Hebrew but that the two languages are not even members of the same language family – English from Indo-European and Hebrew from Afroasiatic.

It seems most likely that spoken language had to be invented many, many times all around the planet. As human groups grew beyond the nuclear family, the social relationships grew more complex and would always be strained by environmental changes. Having a fairly robust medium of communication would provide such an evolutionary advantage that language almost certainly were invented at several different times and places. As local languages interacted and the speakers learned from one another, the languages would change, grow more complex, reflect certain aspects of one another. In time one language could grow strong enough that, like Proto-Indo-European, it would spread and spawn a language family.

**Latin alphabet.** By the time of classical Latin, towards the end of the Roman Empire, the Latin (sometimes called Roman) alphabet consisted of 23 letters, with no <j>, <u>, or <w>. In the 7th century Irish and Roman missionaries in England began to replace the runes of the old Germanic runic alphabet with letters from the Latin, and during and after the 11th century the replacement was extended by the Norman scribes. The <j, u, w> that were missing in the classical Latin alphabet were added to the English alphabet through a fairly complex set of distinctions:

In Old English the letter <i> usually spelled a vowel sound but sometimes it spelled the consonant sound [y], which the Norman scribes spelled with continental <j>. By the 17th century writers distinguished <i> and <j>, which was an elaboration of <i>, with the vowel sounds written <i>
and the consonant sound \( [j] \) written \( <j> \).

Roman \( <v> \) could spell either the vowel sounds we spell with \( <u> \) or the consonant sound \( [v] \), sometimes \( [w] \). By the 17\(^{th} \) century the letters \( <u> \) and \( <v> \) were distinguished with \( <u> \) as usually vowel, \( <v> \) as consonant.

In Old English \( [w] \) was spelled \( <u> \) and \( <uu> \) (thus its modern name “double \( <u> \)”), and sometimes double \( <v> \), \( <vv> \). In the 8\(^{th} \) century the rune wynn \( <\text{Þ}> \) was used to spell \( [w] \), but later Norman scribes replaced wynn with \( <w> \), the ligature – or joining – of the old \( <vv> \).

**Latinization.** During a brief classical revival in the 16\(^{th} \) century some words were respelled to show their Latin sources. For instance, our word *debt* was commonly spelled \( <\text{dette}> \) in Middle English, but a \( <\text{b}> \) was introduced, with no change in pronunciation, to show its Latin source *dēbitum*. Other examples include these: Middle English *enditen* became *Early Modern* *indict* after the Latin *indictare*. Middle English *faute* became *Early Modern* *fault*, with a change in pronunciation, after Latin *fallita*. Middle English *langage* became *language* after Latin *lingua*. In some cases the motivation may not have been so much an enthusiasm for Latin as a distaste for French, the Hundred Years’ War with France having finally ending in the late 15th century: Middle English *columgne*, which came from French, became *column* after Latin *columna*. *Solemn* was the same.

**Lie, lay.** These two verbs have caused trouble for 800 years, and although they are not strictly speaking a spelling problem, they are worth a few words here. *Lie* is an intransitive verb – that is, it does not take a *direct object*; *lay*, which does require a direct object, is transitive. *Lie* means basically “to recline,” as in *I lie on the couch every Sunday until noon.* *Lay*, on the other hand, means “to place something down, to cause it to lie,” as in *I lay the mail on the hall table every afternoon.* Beyond their similarity of meaning, their inflection reveals another source of confusion:

<table>
<thead>
<tr>
<th></th>
<th>Lie</th>
<th>Lay</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(^{st}) &amp; 2(^{nd}) person singular, present tense</td>
<td><em>lie</em></td>
<td><em>lay</em></td>
</tr>
<tr>
<td>3(^{rd}) person singular, present tense</td>
<td><em>lies</em></td>
<td><em>lays</em></td>
</tr>
<tr>
<td>Past tense</td>
<td><em>lay</em></td>
<td><em>laid</em></td>
</tr>
<tr>
<td>Past participle</td>
<td><em>lain</em></td>
<td><em>laid</em></td>
</tr>
<tr>
<td>Present participle</td>
<td><em>lying</em></td>
<td><em>laying</em></td>
</tr>
</tbody>
</table>

The confusion is aggravated by the fact that the 1\(^{st}\) and 2\(^{nd}\) person present of *lay* and the past tense of *lie* are *homonyms*, spelled \( <\text{lay}> \) and pronounced [lā].

**Long vowels.** In Modern English the long vowels are \( [\text{ä}] \) as in *bait*, \( [\text{ē}] \) in
beet, [ɪ] in bite, [o] in boat, [ʊ] in boot and [yʊ] in beaut. Also see Vowel sounds and Great Vowel Shift.

**Lose, loose.** These two are not quite homonyms and not quite homographs. They and their present participles, losing and loosing, can cause problems for spellers. Remember that the past tense of lose is lost, both with a single <o>. And remembering the old expression “loosy-goosy” may help with the <oo> in loose.

**Medal, metal, mettle, meddle.** Medal and meddle are homophones, as are metal and mettle. But because the sound [t] in the middle of words can sound very much like [d], all four of them are near homophones. Medal and meddle are unrelated in meaning or history, but mettle was originally a spelling variant of metal, and over the centuries their spelling and meaning have diverged. Medal and metal are both instances of the French Lemon Rule. Mettle and meddle both contain VCCle strings. The base of medal is med4 “middle, half,” apparently due to an early use of the word to refer to a coin worth half a Roman denarius. The base of meddle is medd1 “mix” and the suffix -le1 marks frequentative verbs.

**Metaphor.** In metaphor a word or phrase is used to refer to something that is categorically different from, but is perceived as being in some important way similar to, what it normally refers to – as when Keats refers to a Grecian urn as a “foster-child of silence and slow time,” encouraging the reader to explore the similarities. But metaphor is not restricted to poetry. It has played a major role in the evolution of English, including English spelling: For instance, it is involved in the often-maligned folk etymology, by which users employ analogy, or metaphor, to make an unusual form more similar to some perceived pattern, as in the following:

<table>
<thead>
<tr>
<th>Original Word</th>
<th>Evolved To</th>
<th>By Analogy With</th>
</tr>
</thead>
<tbody>
<tr>
<td>chaise longue “long chair”</td>
<td>chaise lounge</td>
<td>lounge</td>
</tr>
<tr>
<td>ele, eill “wing of a church”</td>
<td>aisle</td>
<td>French isle and aile “wing”</td>
</tr>
<tr>
<td>foremost (form “first” + -est )</td>
<td>foremost</td>
<td>fore and most</td>
</tr>
<tr>
<td>igland</td>
<td>island</td>
<td>isle</td>
</tr>
<tr>
<td>pentice</td>
<td>penthouse</td>
<td>house</td>
</tr>
<tr>
<td>rime</td>
<td>rhyme</td>
<td>rhythm</td>
</tr>
<tr>
<td>shame-fast</td>
<td>shamefaced</td>
<td>face</td>
</tr>
<tr>
<td>up so doun “up as if down”</td>
<td>upside down</td>
<td>side</td>
</tr>
</tbody>
</table>

But at an even deeper level metaphor can be said to structure how we think and talk about a given topic. For instance, the unconscious metaphor “Love is madness” leads to our use of such language as “I’m crazy about her. She drives me out of my mind. He constantly raves about her. He’s gone mad over her. I’m just wild about Harry. I’m insane about her.” (G.

**Metonymy.** Metonymy and metaphor are two basic methods of thought and thinking. In metaphor a word is used to refer to something that is categorically different from, but is taken as being in some way similar to, what the word normally refers to. So metaphoric relationships are based on similarity. Metonymic relationships are not based on similarity but rather on, broadly, contiguity – that is, on the way things can be associated or brought together – for instance, as part-and-whole, cause-and-effect, more-or-less, first-and-next, thing-and-quality, spatial or temporal adjacency.

The most common metonymy is synecdoche: the identification of a whole thing with one of its parts, as when we speak of several head of cattle or when we speak of the American government as Washington or of a credit card as plastic. In the evolution of a language a common form of metonymy is synecdochic clipping, or abbreviation – as when polysyllables like gasoline and automobile are clipped to gas and auto.

For more on metonymy go to “Metaphor and Metonymy” and “Metonymic Variation” in Orthography as an Evolving Complex System.

**Middle English.** Middle English is English as it was spoken and written from about 1150 to 1500. It was a period of great change in how the language sounded, how sentences were constructed, and how words were spelled. This change was in large part due to English’s interaction with Norman-French and Anglo-French (See Norman scribes, Norman Conquest) and the Anglo-Norse dialects spoken by Vikings and their descendants who settled in eastern England. Due mainly to geo-political reasons, by the 14th century English was largely reestablished as the official language of England. And by the 15th century Anglo-French had all but disappeared, and Standard Written English had begun to develop.

**Minims.** In medieval script a minim (Latin, “least”) was a small downstroke rather like J used to form letters like <u, n, m, i>. Since the scribes tended to crowd letters and words together, uninterrupted strings of minims could produce problems for readers. For instance the word minim itself could become something like •••••••••• – and the <um> in the Old English cuman “come” would contain the string “•••••”. Norman scribes often changed <u>’s to <o>’s to break up problematic strings of minims, thus our modern word come. It was this legibility problem that also led to the dot over the <i>, and by extension the <j>.

**Missionaries.** During and after the 6th century, there were several Christian missions from Ireland to the north of Britain. Over the following centuries the Irish missionaries built abbeys and monasteries throughout Britain, especially in the north. In 597 Pope Gregory sent a mission from
Rome to southern England, led by Augustine, who became the first Archbishop of Canterbury. These Gregorian missionaries were much more active in converting the pagan Britons and Anglo-Saxons than were the Irish missionaries in the north.

The Irish missionaries wrote in an Irish version of the Roman alphabet called the Insular hand. Though it was the script used in many Old English manuscripts, it had little if any lasting impact on the development of English spelling. English spelling was much more influenced by the dialect of Wessex and Alfred the Great, which was based on the Roman alphabet as it had been spread in the south of England by the work of the mission of Augustine.

Misspellings. As any teacher well knows, misspellings come in a myriad of shapes and sizes. Some are really wrong (*sufishuntly) and others are heart-breakingly close (*suffitiently). But with our all-or-nothing approach to spelling, both *sufishuntly and *suffitiently are equally incorrect. Alas, we cannot say things like “You’re getting better. You spelled this word 85% correctly.”

The following are common misspellings taken from A Casebook of Misspellings, with Observations. The misspellings are listed with a leading asterisk:

Some are due to errors in twinning:
• compelling, *compeling
• forgotten, *forgoten
• jarring, *jaring
• patrolling, *patroling
• picnicking, *picnicing The insertion of <k> is a form of twinning.
• shoveling, *shovelling

Some are due to errors in silent final <e> deletion:
• spacing, *spaceing
• spiny, *spiney

Some are due to problems with <i> before <e>:
• freight, *frieght
• receipt, *reciept

Some are due to problems with hard and soft <c> and <g>:
• eligible, *eligable
• intrigue, *intrige
• serviceable, *servicable
• spaghetti, *spagetti

Some are due to over-simplified phonetics, as with the sounds [k] and [n] in
• chrome, *crome
• knack, *nack
• opaque, *opake

Some are due to missing letters and sounds and can be addressed with attention to pronunciation:
arctic, *artic: Arctic explicates to arctic2. It is from Latin arcticus “bear.” The geographical sense is “of the north, of the constellation Ursa Major (the Great Bear) and the north star.” There are accepted pronunciations with and without that first [k], but no variant spellings.

aspirin, *asprin: Aspirin is a tradename that contracts and merges acet and spiraeic acid. aspirin03. There are two accepted pronunciations: [əs'pɛr-ɪn] and [əs'prɪn], but no variant spellings. So remember that middle vowel.

government, *goverment: Government explicates to govern+ment). The [n] is easier to hear in the related govern, governor, and governess.

library, *liberry: Library explicates to libr2+ary2, from Latin liber. The base libr2 “book” also occurs in to librarian and libretto “little book, the text of an opera”, where the first [r] is easier to hear.

twelfth, *twelth: Twelfth twelf+th)1 is obviously related to twelve, with the <f>/<v> pair that we see in words like shelf, shelves; wife, wives; belief, believe. That soft fricative [f] is easy to lose right next to the fricative [th].

Many misspellings can be addressed with explication:

- accessory, *accesory (ad+c1+cess1+ory) with cess1 as in recess, success, abscess.
- amplifier, *ampliphier ampl1+ify/)+i2+er)01 with -ify) as in dignified, gratify, identify, justify, modify, notify, qualified, terrify, testify.
- dismissing, *dismissing (dis+miss3+ing)1 with (dis- as in disagree, disagree, disappoint, discharge, discover, disease.
- maintenance, *maintainance ; main2+tain1, main2+ten2+ance). See Sets.
- orientation, *oreintation, *orentation, ori1+ent)+at)1+ion)1 with ori1 “rising sun, east, place something to face east, place something correctly.”
- paralyzed, *paralized (para1+lyzɛ+ed)1 with lyze as in analyze, dialyze.
- separate, *seperate (se+par03+ate)1 with par03 as in apparel, disparate, disparity, parade, separation.
- discipline, *disipline Discipline explicates to (dis+cip1+l/e)3+ine)3, showing its close relationship to disciple – as well as its wider relationship to such words as anticipate (anti2+cip1+ate)1, municipal muni1+cip1+al)1, participant parti+cip1+ant)1, principal prin+cip1+al)1, and recipe (re+cip1+e)3. It also shows the need for both the <s> and the <c> in its spelling.

And some misspellings involve just plain odd words: *vacume rather than vacuum, *furlow rather than furlough, *ake rather than ache (AES, 357), or *tounge and *tunge rather than tongue: Tongue has been a trouble maker.
since the 13\textsuperscript{th} century: The OED lists these 16 spellings of it:

<table>
<thead>
<tr>
<th></th>
<th>OE</th>
<th>ME</th>
<th>EMnE</th>
<th>MnE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>tunge</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>tunke</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tonke</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tonge</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>tungge</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tongee</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>toungge</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>towng(e)</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tung</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>toung</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>tong</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>tounghe</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>toong</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>toongue</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tongue</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

Notice that \textit{tongue} was spelled <\textit{tunge}> and <\textit{tounge}> for six centuries each, both of which are today considered misspellings. In the 13\textsuperscript{th} century the scribes began replacing <\textit{un}> with <\textit{on}> to avoid a string of four \textit{minims}. It looks as if in the 15\textsuperscript{th} century they replaced \textit{tonge} with \textit{tongue} to \textit{insulate} the <\textit{g}> from the <\textit{e}> to avoid having the final <\textit{ge}> look like a soft <\textit{g}>. For a technical discussion of the kinds of things that went on with early spellings like \textit{tongue}, see \textbf{Standardization in Early English Orthography}.

And finally: \textbf{Okay, but ... Spellings}. As in you’ve spelled a word correctly, but not the word you were trying to spell:

\textbf{calendar, *calendar}: A calender is a rolling machine used in paper making, from Latin \textit{cylindrus} “roller”. Notice the common <\textit{er}> in \textit{calender} and \textit{roller}. \textit{Calendar} is from Latin \textit{calendārium} “account book”, which is from \textit{calends} “first day of the Roman month, when monthly payments were due.” Notice the common <\textit{ar}> in \textit{calendar} and \textit{calendārium}.

\textbf{colonel, *kernel}: \textit{Colonel} alters the earlier \textit{coronel} “the leader of a column”, thus converging a later spelling with an earlier pronunciation. \textit{Kernel} is from OE \textit{cyrnel} “seed”.

\textbf{corpse, *corps}: \textit{Corpse} [kôrps] “a lifeless body” and \textit{corps} [kôr] “a military group” are essentially two spellings of what was once the same word from Latin \textit{corpus} “body”. \textit{Corps} comes via French with French’s usual
silent final consonants; corpse comes directly from Latin with the final consonants pronounced and the final <e> insulating the <s> so it doesn’t look like a plural.

**dining, *dinning:** The word spelled <dinning> is the present participle of the verb *din* din+n2+ing)1; dining is the present participle of the verb *dine* diné1+ing)1 – a classic case of the VCV vs VCC distinction.

**perfect, *prefect:** Prefix (pre+fect ”a chief administrative officer” is a noun; perfect (per1+fect is usually an adjective or verb. Both contain the base *fect* which earlier meant “set, place”, but has come commonly to mean “make, do”. The prefix (pre- means “before, in front”, and prefect comes from a Latin word meaning “to put at the head of”; the prefix (per1- in perfect means “thoroughly, completely”, so something that is perfect is done thoroughly and well.

**psychic, *physic:** Psychic [sīˈkīk] means “not material, mental”; physic [ˈfīk] comes from a Greek word meaning “to grow” which also gave us physics, physical, and physician, all dealing with very material things. Both come from Greek – thus, the rare <ps> from the Greek letter psi <ψ>and the less rare <ph> from Greek phi <φ>. The simplification <ps> is a rare spelling of [s], initial [ps] not coming easily to the English-speaking tongue, though the OED shows two accepted pronunciations of psychic one without [p], one with.

**weather, *wether:** Wether “a male sheep or ram” comes from the IE root *wet2 “year” and is related to veteran and veal. (IE [w] remained [w] in OE but became [v] in Latin.) It occurs in the compound bellwether “the lead sheep of a flock, which had a bell around its neck.” Weather came from IE *wē “blow” and is related to wind, window, wing and via Latin vent, ventilate. Notice also the near homophone whether, which is one of the <wh> interrogative group who, what, where, which, why, when, whether, wether, and whence.

**were, *where and where, *were:** These two near homographs are often confused. Where is part of the interrogative set with <wh>. Were is a past tense form of is:

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Person</td>
<td>I was</td>
<td>we were</td>
</tr>
<tr>
<td>2nd Person</td>
<td>you were</td>
<td>you were</td>
</tr>
<tr>
<td>3rd Person</td>
<td>he, she, it was</td>
<td>they were</td>
</tr>
</tbody>
</table>

**which, *witch and witch, *which:** Remember that which is one of that interrogatie <wh> set. Witch descends from the IE root *weg “strong”,
which also led to our *wicked* and *Wicca* “a pagan religion”. The OE *wicce* “female sorcerer” is the direct source of our *witch*.

*whole, *hole:* PIE [k] regularly became [h] in English so IE *kailo* “whole, uninjured, of good omen” gave us *whole* and the related words *hale, wholesome, health, heal, holy, holiday, Halloween*. On the other hand, *hole* descends from IE *kel1* “to cover, conceal, save,” which also led to the variously related *hell, hall, hold* “of a ship”, *hull, hollow, helmet*. And since PIE [k] regularly remained [k] in Latin, *kailo* also led to *clandestine* “secret, concealed”, *eucalyptus* (from its well-covered bud), *occult*, and *color* “that which covers” with hard <c> – and to *cell, cellar, conceal* in which the <c> later softened. See also Spelling Demons and Doublets Misspelled.

*Mnemonic.* (That <mn> is pronounced [n].) A mnemonic is a device for helping you remember. Probably the best known mnemonic starts out “Thirty days has September . . . .” And the best known mnemonic in spelling starts out “It’s *<i>* before *<e>* . . . .” For some other spelling mnemonics see Capital, capitol; desert, dessert, deserts, accede. The word *mnemonic* is related to a large group of words dealing with the mind: mental, maniac, reminisce, mantra, museum, amnesty. It is from the name Mnemosyne, the Greek goddess of memory, mother of the muses.

*Modern English.* English as it has been spoken and written since roughly 1550 to the present. The evolution from Early Modern English to Modern English was gradual, so the dates given for the two tend to overlap.

*Morphemes and Morphophonemics.* Morphemes are the smallest content-bearing parts of spoken words. They are the spoken language’s equivalents of the written language’s elements.

Four or five centuries after the Norman Conquest and with the development of Standard Written English, English orthography had evolved well past its phonetic origins to a system that the Elizabethan Richard Mulcaster described as ruled jointly by sound, custom, and reason. English spelling is not simply phonetic; it doesn’t relate spellings just to sounds but also to meaningful elements, overriding sound changes. For instance, in *cats* the noun plural suffix -s3 spells the unvoiced [s], but in *dogs* it spells the voiced [z]. If it were not for the morphophonemic feature of English – that is, if English were strictly phonetic – we would have to have two noun plural suffixes – one spelled <s>, the other <z>. Or consider the base *sign*, which shows up with four different pronunciations in *sign* [sɪɡn], *signal* [sɪɡˈnɛl], *design* [dɪˈzɪn], *designate* [dɪˈzɛtɪɡˈnæt]. Or consider the different pronunciations of the base *photo* in *photos* [fəˈtɒʊz], *photography* [fəˈtɒɡrəˈfi], *photographic* [fəˌtɒɡrəˈfɪk]. Morphophonemics basically deals with the changes in sound that elements undergo in different contexts. For more details, see *EDLL*, *OCEL*, and “Post-Alphabetic Literacy” in Orthography as an Evolving Complex System.
Mulcaster, Richard. In 1561 Richard Mulcaster (ca. 1531-1611) became the headmaster of England’s largest school, the Merchant Taylor’s School in London. In 1582 he published his spelling text, The Elementarie. He used a political allegory to argue that a mature spelling system has to do more than spell sounds, for sounds are too changeable and diverse. In the beginning, to paraphrase Mulcaster:

Sound was King, with complete dominion over spelling. But the sounds of speech vary so much from person to person, place to place, and time to time that eventually in the Kingdom of Sound, confusion set in. The unhappy people petitioned their King for relief, and in time they convinced him to have Custom and Reason join him in the rule of spelling. Reluctantly Sound gave up his complete dominion – though he remained the primary member of the triumvirate: Now English spelling would be controlled jointly by "the propriety of Sound," "the smoothing of Custom," and "the consideration of Reason." Sound remains primary, but his vagaries are made more sensitive to convention and word history because of the influence of informed Custom and more regular and ruly because of the influence of orderly Reason.

Mulcaster's very Elizabethan allegory is not just quaint. It reflects quite accurately the early history of our language: Originally the spelling of Old and Middle English was quite phonetic, and Sound was king. During the proliferation of dialects in the Middle English period and with the spread of literacy, this phonetically-based system led to considerable confusion. By Mulcaster’s time the phonetic system was being replaced by one that also took into account the history and meaning of words and the demands for reason and regularity that one finds in any living system. Custom and Reason were given a larger role.

Mulcaster’s language is very close to Modern English: In the sample below the accents over <o> and <e> indicate missing <n>’s, a contraction used to help typesetters justify the right margin. The <ʃ> was a “long <s>” usual in initial and medial position and with the <sh> digraph. I’m not sure about the accent marks in direc´t. Beyond these essentially typesetting differences and with the exception of the relatively minor spelling differences in waie (way), ar (are), dʃcriptiø (description), proprietie (propriety), and joyntlie (jointly), Mulcaster’s spelling is quite modern. (He is, of course, more generous than we in his use of commas):

The right writing of our Engliʃh therefor by waie of dʃcriptiø is, a certain reaʃonable courʃe, to direc´t the pé by ſuch rules, as ar moʃt comformable, to the proprietie of ſound, the cʃideration of reaʃon, & the ſmoothing of cuʃtom ioynlie . . . .

See also Standard Written English, Morphemes and Morphophonemes.
Nasal sounds. Nasals are **consonant sounds** articulated through the nose – [m] as in *mom*, [n] as in *nun*, [ŋ] as in *song*.

Nonce words. The word *nonce* expands *once* and means something like “this particular occasion.” A nonce word is a word made up for a particular, usually temporary, use. Nonce words often make use of established word endings. For instance, the *-nik* in the Russian *sputnik* led to such nonce words as *kaputnik*, *flopnik*, and *stayputnik* to refer to failed attempts to launch. Though most nonce words do not last long – only for the nonce – some do survive and become established in the English lexicon. For instance, Robert Heinlein in *Stranger in a Strange Land* coined *grak* to refer to a certain kind of intuitive knowing, which has caught on. And apparently any number of *-logy* words also started out as nonce words. Many remain so – such as *caneology* “the science of beating someone with a cane” and *dogology* “the science of dogs.” Other apparent nonce *-ology* words that have become established are *Egyptology* and *neonatology*. Probably many or all of the thousands of words and meanings that appear for the first time in Shakespeare’s writings were intended as nonce words.

Norman Conquest. In October 1066 at the Battle of Hastings in southeastern England, Duke William II of Normandy defeated the army of the English king, Harold II, who, along with his two brothers, was killed. On Christmas day 1066 William, now known as William the Conqueror, was crowned king of England. The Normans were *Vikings* who had settled in northern France and developed their version of French, Norman-French. For our interests the most important effect of William’s conquest was that the Norman-French of the victors became the language of the ruling class in England, replacing Old English. It was the language of government, the schools, and of the Norman elite who replaced English land owners throughout the country. The commoners still spoke English, and Latin was the language of the church. Over time Norman-French evolved into Anglo-Norman, while the English of the mostly uneducated commoners evolved and simplified grammatically to something more like what it is today. By the 13th and 14th centuries French became less of a factor in England, but by then Old English had evolved into *Middle English*. See *Norman scribes* and French influence.

Norman scribes. After the Norman Conquest most of the writing was done by scribes who came from Normandy with their continental writing habits. They introduced a number of changes in English spelling: They replaced some Old English spellings with continental ones, and they introduced some spellings rarely, if ever, used in Old English. They replaced the Old English runes yogh <ȝ>, thorn <þ>, and edh <ð>: They replaced yogh with <gh> in words like *right*, *brought*, *through* and with <y> in words like *year*, *ready*, *young*. They replaced thorn and edh with <th>. They replaced the usual Old English spelling of [kw], <cw>, with <qu>, and
Old English <hw> with <wh>. In several Old English words they changed <u> to <o> especially adjacent to <n> and <m> to avoid strings of minims – for instance, in Old English cuman, the source of our word come. They used three letters that were rarely if ever used in Old English: <q, v, z>. To spell [ch] they increased the use of <ch> and <cch>, the latter of which in time became <tch>. They introduced the <sch> spelling of [sh], which in time simplified to <sh>. And they used <k> before letters that spelled front vowel sounds – <e, i, y>.

Norse and Anglo-Norse. Norse refers to the languages of the Viking invaders and others from ancient Norway, Sweden, and Denmark. Anglo-Norse was the dialect that evolved among the Norse settlers who occupied a large part of eastern England beginning in the 9th century. Since Norse and English are Germanic languages, there are many similarities between them, and their interaction influenced English significantly. Signs of the interaction are still with us – for instance, in the pair shirt, which is native Old English, and skirt from Norse. The CommonWords database lists 247 other words from Norse sources.

Nouns. One clue to identifying nouns is to say that a noun refers to a person, place, or thing – to which we might add quality or act. It’s a good clue, but there can be problems figuring out exactly what a thing is – or a quality. And of course there can be some confusion separating an act (as a noun) and an action (as a verb). See Conversion.

A second clue is “A noun is any word that does not end in -‘s) and fits into this slot: “The __________ seemed okay” – though it can be difficult at times to find a setting where such a sentence makes sense, especially with abstract nouns – as in “The freedom seemed okay.”

And a third clue is that several common suffixes regularly form nouns:

-<age) package -<ism) capitalism
-<ance) assistance -<ist)1 violinist
-<ary)2 secretary -<ity) sanity
-<ence) repentence -<ment) government
-<er)01 ruler -<ness) toughness
-<ience) patience -<or)2 professor
-<ion)1 solution -<th)2 growth

Notice how many of these suffixes form abstract nouns that cannot be perceived by the senses – not seen nor heard nor tasted nor touched: -<ance), -<ence), -<ience), -<ism), -<ity), -<ment), -<ness).

Number Words. All of our number words come from Indo-European (IE) roots. The IE root *oi-no meant “one”, leading to Old English (OE) an, which led to today’s indefinite articles an1 and a, and to any. Over time OE an became on (sometimes spelled with an initial <w>), leading to our one, once – and the negative none “not one”. In Latin the word for “one” was
Some Unexpected “One” Words. Originally alone and lone meant “all one”. Eleven comes from OE endleofan “one left over (past ten)”. The measurement words inch and ounce come from Latin words with the sense “one twelfth”. Atone explicates to at1+one: to atone is to make yourself at one with someone. The negative words null, nullify, and annul come from Latin words meaning “not any, not one.” The prefix non- in such words as nonsense, nonconformist, nonperson, and nonviolent in Latin also had the sense “not one.” Finally, and perhaps most surprising of all, nice originally meant “ignorant” – that is “not knowing any.” (Over the centuries nice has gone through an amazing range of meanings: “ignorant; wanton, lascivious; elegant, smart; strange, rare; lazy; effeminate; delicate; over-refined; coy, modest; shy, unwilling; fastidious, hard to please; precise, careful; refined, cultured; hard to apprehend; subtle, small; slender, thin; trivial; full of danger; agreeable, kind.”)

The IE root *dwo meant “two.” PIE [d] consistently became [t] in OE, leading to our two, which up until the 13th century or so was pronounced with [tw]. The root *dwo also led to other [tw] words with the sense “two”: twin, twice; between and betwixt; twelve and twelfth with the sense “two left over”; twenty “twice ten”; and twain “two”.

The Greek word for “two” was dýo, thus our base di1 “two, double” as in dioxide, dilemma “double premise”, diploma “paper folded double”, and the du1 in dual.

The Latin word for “two” was duo, the source of our double. In Old Latin <dvi, dui> sometimes changed to <bi>, as in our bicameral “two chambers”, bicycle, bisexual, binary, and combine “two together”.

Some Unexpected “Two” Words. Four more words in which <tw> carries the sense “two” are twig “forked, split in two”, twilight “two lights”, twine “two strands”, and twist “two things turned together”. Balance comes from a Latin word with the sense “two pans,” like the scales being held by the statue of Justice. Biscuit had the original sense “twice cooked”. In Middle English doubt “to be of two minds” was commonly spelled doute but the <b> was added in the 16th century in imitation of Latin dubitére. Our word dozen, explicates to do3+zen2, in which do3 means “two” and zen2 means “ten.”

The IE root *trei meant “three”. PIE [t] regularly became [th] in OE, leading to our three and other words with the sense “three”: third, thrice “three times”, thirteen “three plus ten”, thirty “three times ten”, triangle “three
Some Unexpected “Three” Words. The Latin word *tribus* referred to one of the three ethnic groups of ancient Rome. *Tribus* is the source of our word *tribe* and *tribute* “something paid to a tribe”. It also led to other <tribute> words, such as *attribute*, *contribute*, *distribute*, and *tributary*. The word *travel* comes from *travail* “suffering”, which originally referred to three stakes in an ancient torture machine. The origin of Latin *testis* “witness” is debated, but probably it meant something like “I stand one of three (between a disputing pair)”. It leads to our *testify* and *testament*. Apparently in time a man’s gonads were taken as witness of his manliness, leading to *testes*, *testicles*, and *testosterone* “male hormone”.

---

The IE root *kweter* “four” was the source of OE *fêower* “four”, though it is not clear how PIE <kw> became OE <f>. Our *four* has the regular expansions *fourteen* “four plus ten” and *fourth* (*four1+th*)1. But it is not clear why *forty* is no longer spelled <forty>, which it often was from the 13th through 18th centuries. *Fortnight* “fourteen nights” contracts the earlier *fourtene night*. PIE [kw] was often spelled <qu> in Greek, leading to *quart* “a fourth of gallon”, *quarter*, and *quartet*.

Some Unexpected “Four” Words. How to *quarter* came to mean “to house, especially in the military, is yet another unclear thing about “four” words. Perhaps it was common to house soldiers in a certain quarter, or fourth part, of a town – no one seems to know. *Quarantine* refers to a period of forty days. *Trapeze*, *trapezium*, and *trapezoid* all come from an old Greek word that referred to a four-legged table. They are related to words with the base *tetra* “four”, as in *tetrameter* “four beats” and *tetrapod* “four-footed”. Our word *square* comes from the Latin *exquadrâre* “to square something” and is the source of *squad*, *squadron* – and with more contractions of sound and changes in spelling, *cadre*. It is also the source of *quarry*, apparently because in a quarry stones are cut into squares.

---

The IE root *penkwe* and its variant form *kwenkwe* meant “five.” PIE [p] regularly became [f] in OE – thus, *five*, *fifty*, *fifth*, *fifteen*. The PIE <p> remained <p> in Greek, which gave us *pentagon*, *pentameter*, and *pentathlon*. *Pentecost*, the 7th Sunday after Easter, is the “fiftieth day”.

In Latin PIE <kw> regularly became <qu>, as in *quintet*, *quintuple* “times five” and *quint2*, short for *quintuplet*. In Great Britain a *quintillion* is the 5th power of a million, but in the United States it is only the 3rd power of a million, defying the etymology – though both are still pretty big numbers.
The <quin> meaning “five” occurs in several long and technical words like quinquagenarian “one who is fifty years old” (quinqu+agen)+arian). The suffix -agen) means “times ten” and the suffix -arian) forms adjectives and nouns.

Some Unexpected “Five” Words. Finger earlier had the sense “one of five” and is closely related to fist, which by way of the Dutch also gave us foist, literally “to take in hand”. The name of the game Keno refers to the set of five winning numbers. The name of the board game Parcheesi is a respelling of a Hindi word meaning “25”. The name of the fruity drink punch comes from another Hindi word meaning “five”, referring to the number of its ingredients.

Six, seven, eight and nine can be treated more briefly. IE *s(w)eks “six’ leads to our six and its forms and to some rare variations. More or less common unexpected “six” words are sextant, a navigational instrument containing an arc that is one-sixth of a circle; siesta from a Latin word referring to the sixth hour of the Roman day, noon; and Sistine, named for Pope Sixtus IV.

IE *septm “seven” leads to our seven and to the base sept “seven” as in septet. The name September refers to the seventh month of the Roman calendar.

IE *oktō(u) “eight” leads to OE eahta via a regular sound change and thus to our eight and its normal expansions. It also leads to Latin octo “eight” – and thus octave “the eighth note above or below a given tone”; octet, octogenarian, and octopus “eight feet”. In the old Roman calendar October was the eighth month.

IE *newn “nine” leads to our nine and its contraction nin in ninth. It also leads to the Latin nov as in novena “nine days of prayer” and November, the ninth month of the Roman calendar. Less obviously it leads to noon “the ninth hour of the day by Roman reckoning”; and ennead “a group of nine”. Brunch combines and contracts breakfast and lunch, the latter of which comes from noon.

The [d] in IE *dekm “10” became in OE [t], thus leading to our ten, teen1 “plus 10”; ty1 “times 10”; and tenth. PIE [d] remained [d] in Latin, leading to such words as decimal, decade, decagon, and December, the tenth month in the old calendar. IE *dekm had a suffixed form *dkm-tom “ten times ten, 100”, which in time contracted to *kmtom, and led to Latin’s centum “100”, originally with a hard <c>. Centum led to the following with soft <c>’s: cent “100th of a dollar”, century, centennial, and percent “per 100”. By Latin the
Some Unexpected “Ten” Words. Decimate originally meant “to kill or destroy one tenth”; a dime “a tenth part”; dicker, probably comes from an earlier “set of ten”; in dozen do3+zen2 the base zen2 means “ten” (and do3 means “two”); a dean was originally a commander of ten; in fortnight the base fort3 contracts an earlier fourteen; and tithe is from an OE word meaning “a tenth”.

Noun plurals. Far and away most nouns form their plurals regularly by adding the suffix -s)3: head, heads; barn, barns; car, cars; tree, trees. After voiceless sounds, -s)3 is pronounced [s] as in cats, but after voiced sounds it is pronounced [z] as in dogs. Most compound nouns form their plurals regularly: assistant directors, babysitters. sit-ins. But, though regular, some can be surprising: notaries public, commanders-in-chief, passersby.

Singular nouns that end with a sibilant sound – that is, [s, sh, z, zh, ch] regularly take the suffix -es)2, pronounced [z]: kiss, kisses; fox, foxes; wish, wishes; buzz, buzzes; church, churches. In nouns like house with its plural houses, the sound of the first <s> voices from from [s] to [z].

Singular nouns that end with <y> preceded by a consonant letter, regularly change the <y> to <i> and add -es)2: cry, cries cr\(y1+i2+es)2. Also flies, spies, tries, etc.

Singular nouns that end with an <o> preceded by a consonant vary between -s)3 and -es)2. Most of them take the regular form, -s)3: pianos, altos, twos, egos. Many take either -s)3 or -es)2: banjos, banjoes; zeros, zeroes; mottos, mottoes; innuendos, innuendoes. With either of these sets it's always safe to use -s)3. But a few nouns that end with <o> preceded by a consonant take only -es)2, the most common being echo, embargo, go, hero, Negro, potato, tomato, and veto. With these eight the plural should be formed with -es)2. (When hero is used in the compound hero sandwich, its plural is often spelled heros.)

Seven nouns from Germanic form their plurals irregularly via ablaut – that is, by changing the internal vowel: foot, feet; goose, geese; louse, lice; man, men; mouse, mice; tooth, teeth; woman, women. Notice that in woman, women though the <a> changes to <e> their pronunciation is the same, but the pronunciation of the <o> changes from [u\(a\)] in woman to [\(a\)] in women.

Even fewer plural nouns use the Germanic suffix -en)05 or its expansion -ren): brethren, menschen, oxen – and children, which blends two Old English plural suffixes -ru) and -en)05.

A few nouns – elf, half, self, shelf, wolf – change the final unvoiced [f] to voiced [v], change the spelling, and add -es)2: elf, elves; half, halves; self, selves; shelf, shelves; wolf, wolves.

Some foreign nouns retain their foreign plurals: alga, algae; alumnus, alumni; analysis, analyses; criterion, criteria; larva, larvae; stimulus, stimuli.

Some foreign nouns have both foreign and regular plurals, the foreign
plurals being more common in technical registers: antenna, antennae, antennas; cherub, cherubim, cherubs; index, indices, indexes; minimum, minima, minimums; tempo, tempi, tempos; vertebra, vertebrae, verteb ras.

Some nouns do not change to show plural: deer, elk, fish, moose, sheep. And most abstract nouns do not have plurals.

The plurals of letters and numbers use ‘s: three r’s and five 4’s.

Though the huge majority of nouns take the regular noun suffixes -s)3 or -es)2, when in doubt, check a good dictionary.

Old English, or Anglo-Saxon. Old English is English as it was spoken and written from roughly the middle of the 5th century until the mid 12th century by which time it had evolved into Middle English. It developed in England among the Anglo-Saxon settlers – mainly the Angles, Saxons, Jutes, and Frisians from northern Europe. It is a Germanic language. Its early evolution was much influenced by the language of the Vikings and other Norse settlers and later by Norman-French, the language of the Norman conquerors. Old English literary works first appear in the 7th century. Old English is the language of Beowulf. After the Norman Conquest in 1066 Old English was replaced by Norman-French as the language of power and schooling in England.

Old English was a heavily inflected language. For instance, the Old English version of our definite article the had 23 different inflected forms. Singular nouns were inflected for four cases – nominative (subject), accusative (direct object), dative (indirect object), and genitive (possessive). Pronouns were inflected for singular, plural, and dual numbers and for masculine, feminine, and neuter genders. By late Middle English, after centuries of English being the language of the unschooled, this complex system had simplified to basically what we have today, with less reliance on inflections and more on word order.

Open and closed syllables. Open syllables end with a vowel sound, closed syllables with a consonant. Thus, in open, the first syllable, [ô], is open; the second syllable, [pên], is closed. See Syllables.

Oxford English Dictionary (OED). The OED is a multivolume dictionary that describes the history of English words as their senses and spellings changed from Old to Middle to Early Modern to Modern English. Work on it began in 1857, and until the late 1920’s it was published as a series of short unbound fascicles. The first complete edition was published in 1928. There have been several supplements published since 1933, including a four-volume supplement published from 1972-1986. The second edition was published in 1989. In 2000 work began on a third edition, which is projected to be published in 2037. In addition to the print versions, the second edition is available on CD-ROM and on-line at oed.com. The on-line version is upgraded every three months with new material prepared for the third edition. In all of its manifestations the OED is the essential guide to the lives of English words as they enter and sometimes leave the
Palatal sounds and their spellings. Palatals are consonant sounds articulated with the tongue against the palate, or roof of the mouth – [j, ch, zh, sh, y].

[j] In CommonWords [j] is spelled either <j> or <g> about 90% of the time, and it is spelled <g> almost three times as often as it is spelled <j>. It is spelled
- <j> usually in element-initial position (jealous, jockey, rejoice (re+joice, sojourn (so1+journ), rarely in medial position (banjo, cajole, pajamas, pejorative), and in final position only in raj;
- <g> only before <e, i, y> (genuine, magic, clergy, squeegee, origin);
- <dge> only in final position after short vowels (badge, fudge, lodge, midge, pledge);
- <dj> only after (ad- (adjective, adjourn, adjust, adjutant);
- <gg> only in exaggerate;
- <d> only before <u>: gradual, individual, pendulum, schedule. AES, pp. 417-21.

[ch] About two-thirds of the time in CommonWords [ch] is spelled either <ch> or with the doublet equivalent <tch>. It is spelled
- <ch> in any position in the word (achieve, cheese, church, duchess, launch, merchant, orchard, spinach, treachery);
- <tch> always after short vowels, usually in word-final position (batch, etch, hutch, notch, pitch, thatch), but sometimes word-medial (kitchen, ratchet, satchel);
- <c> in Italian adoptions (cello, vermicelli, concerto);
- <tsch> in German adoptions (kitsch, putsch);
- <t> preceding unstressed <u> or <ion> (actual, combustion, digestion, fortune, mortuary, suggestion, virtual). AES, pp. 412-17

[sh] In CommonWords [sh] is spelled <sh> or <t> about 85% of the time, and it’s spelled <t> almost twice as often as <sh>. The following lists the nine spellings of [sh] in descending order:
- <t> preceding an unstressed <i> which is followed by a second unstressed vowel letter, commonly <ion> (abbreviation, description, impatience, initial, notion, ratio);
- <sh> nearly always in initial or final position (bashful, fresh, mushroom, nourish, rubbish, shoulder, shush);
- <c> preceding an unstressed <i> followed by a second unstressed vowel letter (ancient, commercial, glacier, specialty, species, suspicion);
- <ss> preceding <ion> or <u> (assurance, commission, confession, expression, issue, pressure, session);
- <s> preceding <ion> or <u> (dimension, expansion, insure, mansion, sugar, sure, suspension);
- <ch> usually in fairly recent French adoptions (brochure, chagrin,
Chicago, champagne, machine, mustache, parachute) and pistachio (from Italian);
• <sc> especially in the base sci1 “know” (conscience, conscious, omniscience, unconscious);
• <sch> especially in words from German and Yiddish (schlemiel, schmo, schnauzer, schnook, schwa);
• <x> apparently only in sexual. AES, pp. 407-12

[y] In Old English until the 13th century [y] was spelled with the rune yogh, <ʒ>, then with <y>. It is spelled <y> about half the time:
• <y> in element-initial position (beyond, yacht, yellow, youth, yule);
• <j> in fjord only;
• <i> about half the time in the [y]-glide, unstressed and followed by a second unstressed vowel (alien, billion, civilian, companion, dahlia, onion, spaniel);
• <y> in the [y]-glide, in lawyer and sawyer. AES, pp. 459-60

[zh] In CommonWords [zh] is spelled <s> about 85% of the time; the rest of time it’s spelled <g>:
• <s> preceding either unstressed <i> followed by another unstressed vowel letter or <u> (casual, collision, enclosure, erosion, excursion, leisure, measure, usual, visual);
• <g> in French adoptions, where it sometimes has a variant pronunciation with [j] (camouflage, garage, gendarme, mirage, prestige, regime, rouge);
• <z> preceding <u> in seizure and azure;
• <t> in equation only. AES, pp. 421-22

Palatalization. When a consonant sound is palatalized, it is pronounced back in the roof of the mouth, against the hard palate. For instance, the sound spelled <t> in native, [t], is not palatalized; it is an alveolar, pronounced forward in the mouth. But the sound spelled <t> in nation, [ʃ], is palatalized; it is a palatal, pronounced well back against the palate. There is a palatalized spelling when a letter like <t>, which normally spells a nonpalatal sound as it does in native, spells one that has been palatalized, as it does in nation.

For the several different palatalizations see Palatal sounds and their spellings. For more on palatalization see AES, Chapter 30, “The Palatal Sibilants,” pp. 407-22, and “Palatalization” in Spelling for Learning.

Participles. In English there are two participle verb forms: the past participle, usually spelled with the suffix -ed)01, sometimes with -en)02 or -n)1, and sometimes with just a change of vowel, as in have missed, have broken, have drawn, have sung; and the present participle, sometimes called gerunds, always spelled with -ing)1, as in is hurrying. Past participles carry the sense “action completed” and usually take some form of have or be as auxiliary (or “helping”) verbs. Present participles carry the
sense “action still going on” and take some form of to be – as in “The cake was baked” vs. “The cake is baking.” The word participle means “that which shares in the properties of two or more others,” and our past and present participles share in the properties of verb, adjective, and noun – as in “He has broken his arm twice” (verb), “His broken arm was set at the ER” (adjective), and “Breaking his arm was really a stupid thing to do” (noun), and in “The news is breaking” (verb), “The breaking news” (adjective), “Breaking the news was a mistake” (noun).

**Particles.** Particles contribute no content to a word, though they can serve various functions. Particles enter in when elements link together, or concatenate, to form words. The most common particles are the second consonants that are inserted in twinning when a word like run takes a suffix that starts with a vowel: run1+n2+ing)1 = running, the particle <n2> serving the orthographic function of marking the preceding short vowel. Similar particles occur in the assimilation of prefix-final consonants – for instance, the <s1> in assimilation (ad+s1+simil+atê)1+ion)1, where the particle eases pronunciation. Other particles function as linkers, like the letter <o> common in technical words, especially from Greek: ileostomy (ile1+o4+stom+y)3 and ozonosphere (oz+onê)3+o4+sphere. For more on particles see On Explication. A list of all 39 identified particles appears in the **Particles** data table of the Lexis database.

**Partridge, Eric.** Eric Partidge (1894-1979) was born in New Zealand and lived in Australia and England. He was a prolific lexicographer of the English language, having written several dictionaries of slang and etymology. His *Origins: A Short Etymological Dictionary of Modern English* is particularly useful since it gathers together cognate words and explores their sources, senses, and relationships.

**Pastor, pasture.** Although in their day to day use these words may not seem to be related in history and meaning, they actually are: Both are thought to come from an Indo-European root that meant “shelter, feed.” *Pastor* explicates to past6+or)2, the base past6 meaning “shelter, protect” and the suffix -or)2 meaning “one that does.” *Pasture* explicates to past2+ure), the base past2 meaning “feed, food” and the suffix -ure) marking a noun of result or source. There’s a subtle but important difference in the pronunciation of the two: pastor is [p ås'ter], pasture is [p ås' cher]. For that <t> spelling of [ch] in pasture see **Palatalization.**

**(Per1-, (pre-.)** These two prefixes are close enough in spelling and sometimes in pronunciation that it can be easy to transpose the <e> and the <r>. The more common prefix (pre- means “before, prior to, in front of” – as in preamble, precaution, precedent, predict, preface, prefer, prefix, pregnant, prejudice, preliminary, prepare, preposition, prescription, present, preserve, president, pretend, pretense, prevalent, prevent, previously, represent. The less common (per1- is an intensive and means “thoroughly,
“completely,” though its meaning can be hard to detect in most words: perceive, perfect, perform, perfume, perimeter, period, perish, permanent, permission, permit, perpetual, persist, perspective, persuade.

**Periphrastic.** Periphrastic (peri+phrast+ic)\(^1\) refers to the use an additional word in place of an inflectional suffix to indicate possession -- as in “a friend of Bob” rather than “Bob’s friend” -- which often can be doubled up as “a friend of Bob’s.” The prefix (peri+) means “around”; the base phrast meant “to think,” but later came to mean “to point out, say” -- and is a bound form of the free base phrase. Notice that we still use the notion of “talking around” a subject to describe not coming directly to the point.

Another common periphrastic form is the use of more and most to form the comparatives and superlatives of many adjectives and adverbs rather than using the inflectional suffixes +er and +est -- as in “more wise” and “most wise” rather than wiser and wisest.

**Personal, personnel.** Personnel is essentially the French version of personal, which explicates to person+al\(^1\); personnel explicates to personn+el\(^1\). The base person earlier meant “mask, actor,” and personn is a French expansion of person. A French professor once told me that the <nn> in words like personnel, questionnaire, legionario, and mayonnaise indicated that the preceding vowel was nasalized, or articulated through the nose -- a distinction more important to French than English. The suffixes -al\(^1\) and -el\(^1\) both are used to form adjectives and nouns, -el\(^1\) again being more typically French, and receiving more stress than the -al\(^1\) in personal.

**Phonemics vs. Phonetics.** Phonemics deals with phonemes -- that is, with abstract categories of speech sounds, categories whose members are treated as being equivalent in function. Phonetics deals with the concrete and specific speech sounds, or phones, that are members of those phonemic categories. The contrast between phonemics and phonetics parallels the contrast between content and meaning in that content sets the boundaries for using a word to make meanings, and the phoneme sets the boundaries for using phones to do the same. It also parallels the contrast between code and performance, in that code sets the boundaries in which a performance is meaningful.

The symbols for phonemes are conventionally enclosed in diagonals, while symbols of phones are conventionally enclosed in square brackets -- thus, /t/ for the abstract phoneme, and [t] for the concrete phone. (But see the NB at IPA for usage at this site.) Phonetic transcriptions can show added details about phones. For instance, if you hold the back of your hand close to your mouth as you pronounce, say, tick and stick, you should feel a little puff of air with the [t] in tick but no puff with the [t] in stick. That puff of air indicates that in tick the [t] is aspirated, spoken with a puff of air -- but in stick it is not. In phonetic transcriptions aspiration is symbolized with a
Just as a phoneme is realized via phones, some scholars speak of **graphemes** as the written counterpart of phonemes, being realized by **graphs** – or written letters – the written versions of phones. An abstract grapheme is realized via a potentially infinite number of **allophones** – for instance, upper vs. lower case letters or printed vs. cursive or italic vs. roman – and more broadly different letter forms in individuals’ handwriting. Graphemes and their graphs are enclosed in angle brackets – thus, <t> the letter vs. /t/ the phoneme vs. [t] the phone. *OCEL*, p. 452, *EDLL*, p. 161.

**Phonetic spelling.** Phonetic spelling is spelling with a phonetic alphabet in which each speech sound is represented by a single letter and each letter represents a single speech sound. A common example of phonetic spelling is the pronunciation spellings given in dictionaries right after the main entry word. The most ambitious phonetic alphabet is the **International Phonetic Alphabet** (IPA) used by linguists and anthropologists. For centuries people have argued that English spelling should be more phonetic, but see *Morphemes and Morphophonemics* and Richard Mulcaster. For a short but lethal critique of spelling reform, see Jacques Barzun’s “The Naiveté of Spelling Reform” in his *A Word or Two Before You Go . . .* (Middletown: Wesleyan UP, 1986), pp. 92-98.

**Plurals, regular and nonregular.** See *Noun plurals*.

**Possessives.** The possessive case is sometimes called the genitive. In nouns it is usually marked with -’s): the foreman’s sister. Less often it is marked periphrastically with the preposition of: the sister of the foreman. As its name suggests, the possessive case usually indicates possession, but it can indicate a few other similar relationships: Description: children’s toys; agent or doer of the act: the principal’s permission, the doctor’s arrival; the recipient of the act: my neighbor’s murderer. Possessive pronouns do not have apostrophes: his, hers, its, ours, yours.

**Prefixes.** The most clear-cut prefixes carry a prepositional or a negative meaning. Among the prepositionals showing direction or location are (anti- “opposite, against,” (ad- “to, toward,” (epi- “on, over, close to,” (ex- “outside, not, without,” (in-2 “in”, (ob- “inverse,” (pre- “earlier, before, anterior.” Among the negatives are (a-1 “without, not,” (in-1, (non-, (un-. The distinction between prefixes and bound bases is a fuzzy one: Many elements are classified as prefixes in some dictionaries and as bound bases in others. See also “Definition of Prefixes” in the *Introduction to the Lexis Database*.

---

superscript <h>. Thus [tʰɪk] for *tick* with aspiration but [strɪk] for *stick* with none. You can feel similar contrasts in pairs like *pit* and *spit*; *kit* and *skit*, *trip* and *strip*. Such aspirated and unaspirated pairs are allophones – that is, “other, or different, sounds” – within the /t/ phoneme. See IPA for more on such phonetic features. Also *OCEL*, pp. 769-71, *EDLL*, pp. 298-99.
Prepositions. Prepositions come in several sizes and shapes and serve several functions, but we will concentrate on just two types: (1) Most prepositions indicate direction or location: *up, down; to, toward; at, beside; from, away from; on, onto; off, off of; into, in; out of* – and the very general *of*. (2) Some prepositions serve as intensifiers, especially *up* and *down*: *burn, burn up, burn down.* (EDLL, p. 312; OCEL, pp. 801-03)

**Principle, principal.** These two homophones are often confused. *Principle* is a noun that means “truth, law, standard.” *Principal* as an adjective means “main, most important,” and as a noun it means “chief person,” as in the principal of a school. Etymology is not much help in distinguishing the two because they both descend from a Latin word meaning “leader, emperor.” And word structure is not much help either, since their two suffixes, -(e)2 and -al)1, can both form adjectives and nouns. Corny as it may seem, the one help I know is the mnemonic phrase “The principal is our pal” – which at least helps with the youngsters’ most common use of principal. And it may help to know that *principle* has a derived adjective, *principled*, but *principal* does not.

**Principle of Preferred Regularity.** Faced with variant spellings – such as, for instance, *busing* and *bussing* – it makes sense to choose the variant that is most regular – that is, most agrees with the patterns and rules of English spelling. Thus the Principle of Preferred Regularity would urge *bussing* over *busing* since *bussing* is a regular instance of the twinning rule while *busing* contradicts that rule. Choosing the more regular spelling helps the spelling system evolve to greater regularity and simplicity. (This example is complicated somewhat by the fact that we also have the words *buss* and *bussing* that mean “kiss” and “kissing”!) AES, pp. 25-26.

**Procedures.** The word *procedure* refers to things that must happen when elements are linked together. CommonWords recognizes four procedures: simple addition, twinning, silent final <e> deletion, <y>-to-<i> and <i>-to-<y> replacement.

**Processes.** The word *process* refers to changes in pronunciation and sometimes spelling that occurred long ago and have complicated modern English spelling. CommonWords recognize three processes: assimilation, palatalization, and vowel shortening due to suffix rules.

**Quite, quiet, quit.** These three all come from a Latin word that meant “to rest,” a relationship clear enough in *quiet and quit*, less so with *quite*. The details are not settled, but it appears that the progression of senses with *quite* was from “rest” to “free” to “totally, completely (done).”

**[r]-Colored Vowels.** The sound [r] can have a strong effect on the sound spelled by the vowel letter preceding it. For instance, in *mate* the silent final <e> marks a long <a>, [ä], but in *mare*, with <ar>, the vowel is closer to [ê]
than [ä]. On the other hand, the silent final <e> in mare is serving some
diacritical function, as is revealed in the contrast between mare and mar, or
bare and bar, or care and car, fare and far, etc.

The effect of the [r] on the preceding vowel varies in its strength, the mare
vs. mate contrast being among the strongest. The list below shows some of
those effects in vowel spellings and pronunciations. The CommonWords
sound symbols are those used in the CommonWords database. The IPA
symbols are taken from Kenyon, John S., and Thomas Knott. *A
Merriam, 1944:

<table>
<thead>
<tr>
<th>CommonWords sound symbols</th>
<th>IPA</th>
<th>Spellings</th>
</tr>
</thead>
<tbody>
<tr>
<td>[a1r]</td>
<td>/ær/</td>
<td>tariff, carriage, guarantee</td>
</tr>
<tr>
<td>[a3r]</td>
<td>/ɛr/</td>
<td>rare, pair, prayer, wear, heir, where</td>
</tr>
<tr>
<td>[e1r]</td>
<td>/ɛr/</td>
<td>dictionary, berry, bury, cemetery, sheriff</td>
</tr>
<tr>
<td>[e3r]</td>
<td>/ɜr/</td>
<td>dear, peer, weird, pier</td>
</tr>
<tr>
<td>[i1r]</td>
<td>/ɪr/</td>
<td>irritate, tyranny</td>
</tr>
<tr>
<td>[i2r]</td>
<td>/aɪr/</td>
<td>fiery, desire</td>
</tr>
<tr>
<td>[o3r]</td>
<td>/ɑr/</td>
<td>car, barn, park, starred, heart, seargeant</td>
</tr>
<tr>
<td>[o5r]</td>
<td>/ɔr/</td>
<td>war, reward, warrant, dinosaur, nor, roar, floor, source</td>
</tr>
<tr>
<td>[u1r]</td>
<td>/ʊr/</td>
<td>fur, burn, burrow, first, squirrel, term, word, worry, courage, earl, myrrh, colonel</td>
</tr>
<tr>
<td>[u3r]</td>
<td>/ɜr/</td>
<td>amateur, poor, tour, fluorocarbon, during</td>
</tr>
<tr>
<td>[u4r]</td>
<td>/ər/</td>
<td>coward, mother, admiral, soldier, doctor, injury, martyr, cupboard</td>
</tr>
<tr>
<td>[yu3r]</td>
<td>/jɜr/</td>
<td>Europe, curiosity, puritan</td>
</tr>
</tbody>
</table>

For more discussion of [r]-coloring see AES, chapter 25, “Vowels Before

**Recent, resent.** Perhaps the first thing to notice about these two near
homographs is that in recent the stress is on the first syllable, while in
resent it is on the second. Another thing to notice is that recent is an
adjective, while resent is a verb. Recent explicates to (re+cent4; resent to
(re+sent2). The base cent4 means “fresh, new, young” and apparently
appears only in resent and its derivatives. The base sent2 “feel” occurs in
the verbs assent, consent, dissent, and more remotely in sentence,
sentinel, and sentry. In both recent and resent the prefix (re- is an
intensive, so recent means something like “very new,” while resent originally
meant something like “to feel strongly,” but by the 17th century it had picked
up the negative meaning that it has now.

**Reduced vowels.** Reduced vowels are unstressed vowels that have lost
their stressed vowel sound and usually have reduced down to schwa – a
soft “uh” sound like that spelled <a> in *alone* [əˈlōn]. Some vowels reduce down to an unstressed [ɪ] like that spelled <a> in *leverage* [ˈlɛvərɪdʒ] or [ˈlɛvərɪdʒ]. Merriam-Webster dictionaries use a dotted schwa [ə] to show a range of reduced vowel sounds from [ə] to [ɪ]. Reduced vowels can occur in either open or closed syllables.

**Reduplication.** In reduplication a word or part of a word is repeated exactly or with a slight change. Repetition of the whole word is common in words mimicking baby talk: *bye-bye, choo-choo, night-night, no-no, pee-pee, poo-poo.* Rhyming reduplication is a common method of intensification: *superduper, huggermugger, chugalug, peewee, pellmell, hotchpotch.* Reduplication in which a word is repeated with a change of vowel is usually also an intensifier, as in *chitchat, gewgaw, knickknack, lippy-loppity, mishmash, pitapat, rickrack, riprap, seesaw, slipslop, ticktack, zigzag.* *Schm-reduplication* was adopted from Yiddish and is usually derogatory or comic: *fancy-schmancy, smartsy-schmartsy,* and any number of others.

**Register.** A register is a variety of English based on the situation or field in which it is used. A register involves the vocabulary of the field but also the way sentences are constructed and the tone in which they are used. The list of registers would include such things as scientific-technical, religious, legal, medical. But there are also registers dealing with cuisine, education, sports, etc. We could even include a youth register, based on a vocabulary of rapidly changing slang and special constructions, such as “He’s a great guy. Not.”

**Romance Languages.** Romance languages are languages that descended from Latin. The major four are French, Spanish, Italian, and Portuguese. Less well known are Catalan (spoken in Catalonia), Provencal (spoken in southeastern France), Rhaeto-Romance (spoken in areas around the Alps), and Romanian (spoken in Romania). See *Indo-European Languages* and *The Indo-European Family* at Lineages from Indo-European Roots.

**Runes.** Runes are the letters from a group of *Germanic* alphabets in use in northern Europe and probably based on the Latin alphabet. The runic alphabet of most interest to us is the *futhorc,* used by the *Anglo-Saxons,* only four of whose runes are relevant to our discussion: *thorn* <þ> “thorn”; *edh* <ð>; *wynn* <ƿ> “mirth, joy”; and *yogh* <ȝ>, an Old English form of <g>, from the rune *gyfu* “gift”. Thorn and edh were used interchangeably to represent either [θ] or [ð], as in our *thin* and *then,* and were replaced by <th>; wynn was replaced by <w>, and yogh by <g> and <y>. See *Norman scribes.*

**Schwa.** Schwa is a reduced vowel and the most common vowel sound in English. It is the sound you hear at the beginning of the word *alone,* a soft "uh." It is written phonetically with an upside-down <e>: [ə]. Schwa sounds
like [ʊ], except that [ʊ] is always stressed, but schwa is always unstressed, so the word *above* [ə-bʌv] contains both schwa and [ʊ]. (The Merriam-Webster dictionaries use a schwa to symbolize both the schwa sound and stressed short <u>.) It is the most common vowel sound in English because most vowels, when they are unstressed, reduce down to schwa. Schwa can be hard to spell because it can be spelled by any vowel letter and nearly any combination of vowel letters. In the following words the bold large letters are twenty different spellings of schwa:

<table>
<thead>
<tr>
<th>abridge</th>
<th>sergeant</th>
<th>cabinet</th>
<th>obscure</th>
</tr>
</thead>
<tbody>
<tr>
<td>hallelujah</td>
<td>bureau</td>
<td>parliament</td>
<td>tortoise</td>
</tr>
<tr>
<td>mountain</td>
<td>sovereign</td>
<td>patient</td>
<td>miraculous</td>
</tr>
<tr>
<td>epaulet</td>
<td>luncheon</td>
<td>anxious</td>
<td>calculus</td>
</tr>
<tr>
<td>emergency</td>
<td>righteous</td>
<td>nasturtium</td>
<td>oxygen</td>
</tr>
</tbody>
</table>

Old English had a complex system of inflectional suffixes, but in the evolution from Old to Middle English, many of these suffixes simplified to a neutral schwa sound. This vestige of old inflectional endings came to be spelled <e>. In Middle English (from about the 12th through the 15th century) final <e> was pronounced as a separate weak syllable, the schwa [ə]. Over the years the final <e> fell silent, though it tended to stay in the spelling. See Silent final <e> functions and deletion and Vowel Lengthening and Shortening.

Semantics and semiotics. See Content and meaning.

Shakespeare, William. William Shakespeare was the greatest poet and playwright of the Early Modern English period, a period of many great writers. He is widely said to be the greatest writer of the English language. He is certainly the most quoted single author in the OED, with 33,051 citations, nearly double the number for the second place writer, Sir Walter Scott. Beyond that, according to the OED, 1,498 of his quotations are the first evidence for a word in English, and 7,613 are the first evidence for a particular meaning of a word. He clearly had an astonishing effect on the development of the English lexicon – and to a certain extent, its spelling. Though at times difficult, Shakespeare’s English is quite accessible to modern readers, as in these opening lines from *Macbeth*. The <ʃ> is the “long <s>” used in place of <s> in non-final positions:

Thunder and Lightning. Enter three witches.

1. When shall we three meet againe?
   In Thunder, Lightning, or in Raine?
2. When the Hurley-Burley’s done,
   When the Battale’s loft, and wonne.
3. That will be ere the jet of Sunne.

Shortening rules. In English there is a small group of rules that account for the presence of short vowels in strings where we would normally expect long ones. See, for instance, the **Third syllable (vowel) rule**; **Suffix rules**, and **French Lemon (or Stress Frontshift) Rule**. See also **Vowel lengthening and shortening**.

Short Word Rule. English spelling restricts two-letter words – such as *be*, *by*, *do*, *is*, *to*, *we*, and *an* – to a small group of function words and very common verbs. That's one reason we have some otherwise unnecessary final double-consonants – such as in *egg*, *ebb*, *add*, and *err*. Also some silent final <e>'s appear at the ends of words to keep them longer than two letters and thus distinct from words in that select two-letter group: *be* vs. *bee*; *by* vs. *bye*; *do* vs. *doe*; *to* vs. *toe*; *we* vs. *wee*. This usage extends to otherwise two-letter words in general: *dye*, *eye*, *foe*, *pie*, *rye*, *tee*, *vee*, *woe*, *zee*. See AES, pp. 87-89.

**Shudder, shutter.** *Shudder* is a verb that explicates to shudd+er)03; the base *shudd* means “shake, tremble,” and the suffix -er)03 marks frequentative verbs. So to shudder is to shake or tremble repeatedly. *Shutter*, on the other hand, is a noun that explicates to shut+t2+er)01; the base *shut* means “shut,” and the suffix -er)01 means “one that does,” so a shutter is something that shuts or closes. In spite of all these differences between them in meaning, function and form, the two words are easy to confuse because they are also near homophones: In *shutter* the <tt> spells [t], but when the sound [t] comes between two vowels, especially if the first vowel is stressed, it can sound very much like [d]. Linguists call it a flapped [t]. The same problem exists with pairs like *ladder*, *latter*; *pedal*, *petal* and *hardy*, *hearty* – and *medal*, *metal*, *mettle*, *meddle*.

Silent final <e> functions and deletion. Silent final <e> can serve five different functions:

• **marking long vowels**: A silent final <e> regularly will mark a stressed vowel in front of it if there is only one consonant letter between them:

<table>
<thead>
<tr>
<th>Short Vowel Sounds</th>
<th>Long Vowel Sounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ã] cap</td>
<td>[ã] cape</td>
</tr>
<tr>
<td>[ĕ] met</td>
<td>[ĕ] mete</td>
</tr>
<tr>
<td>[ĭ] gyp, fin</td>
<td>[ĭ] type, fine</td>
</tr>
<tr>
<td>[ö] not</td>
<td>[ö] note</td>
</tr>
<tr>
<td>[û] dud</td>
<td>[û] dude</td>
</tr>
</tbody>
</table>
There are three holdouts to this rule: In words that end in <aste>, <ange>,
and <le> the silent final <e> will usually mark the preceding vowel as long
even though there are two consonants between the vowel and the <e>:
paste vs. past; range vs. rang; rifle vs. riffle.

- **marking soft <c> and <g>**: Whether or not it affects the preceding
vowel, sometimes silent final <e> affects the preceding consonant –
namely, the sounds spelled by <c> and <g>: pac vs. pace, lac vs. lace,
mac vs. mace and dog vs. doge, hug vs. huge, sag vs. sage, stag vs. stage,
wag vs. wage. See [Soft and hard <c> and <g>].

- **marking voiced <th>**: A silent final <e> will mark a preceding <th> as
voiced, pronounced [th]: bath with voiceless [th] vs. bathe with voiced [th].
Also cloth vs. clothe, teeth vs. teethe, loath vs. loathe, breath vs. breathe.
See [Voiced vs. voicless (unvoiced) sounds].

- **insulating word-final <v> and <u>**: Long ago <v> and <u> were
different forms of the same letter. See [Latin alphabet]. For reasons that go
back hundreds of years, we use silent final <e> to insulate an otherwise
word-final <v>: curve, give, groove, have, shelve, and thief. And except
for a few recent foreign borrowings – like gnu, bayou, and tabu – we also
use it to insulate an otherwise word-final <u>, as in plaque, tongue, league,
statuesque.

- **insulating base-final <s> and <z>**: Silent final <e> is sometimes used to
keep a word from ending with a base-final single <s> or <z>: For instance,
in tens ten+s and laps lap1+s the <s> is the plural suffix; in tense and
lapse the <s> is part of the base, and the silent final <e> keeps them from
looking like the plural nouns tens and laps. If the <s> has a short vowel
right in front of it, another <s> will be added rather than a silent final <e>:
mass, mess, miss, moss, muss. The same pattern holds for the letter <z>.
The letter <z> is fairly rare in English, and the sound [z] is most often
spelled <s>. As with <s>, we tend to avoid ending a word with a single <z>.
If there is a short vowel preceding the <z>, we add a second <z>, as in
fuzz, fizz, and jazz, but notice the recent adoption fez. If there is a
consonant or a long vowel preceding the <z>, we add a silent final <e>, as
in bronze, wheeze, and booze.

- **enforcing the Short Word Rule**.
Some silent final <e>’s are [fossils] with no modern function.

Deleting silent final <e> is fairly straightforward: We delete a silent
final <e> that is marking a soft <c> or <g> whenever we add a suffix that
starts with <e>, <i>, or <y>: lacé+ed]1, lacé+ing]1, lacé+y]1 and except for
some local cases involving stems ending <ee>, <ie> and <oe>, we delete
any other silent final <e> when we add a suffix that starts with any vowel letter.

Those local cases with <ee>, <ie>, and <oe> are these:
- With stems that end <ee>, we delete the final <e> only when adding
suffixes that start with <e>: teé+ed]1 but teé+ing]1.
- With stems that end <ie>, we delete the final <e> when adding a
suffix that starts with any vowel except <i>, in which case we change the
<i> to <y> and then delete the <e>: died, but dying.

- With stems that end <oe>, we delete the final <e> when adding suffixes that start with any vowel except <i>: toed but toeing.

AES, pp. 145-60.

**Silent letters.** In one sense all letters are silent since they are things you see not hear. But in another sense, a letter is silent if it is there in the written word but you can hear no trace of it when the word is spoken. In that sense the <b> in climb can be said to be silent, as can the <p> in cupboard or the [b] in subpoena. The sounds [m, b, p] are all bilabial consonant sounds, and when two bilabials come together there is a tendency for one of them to get lost in pronunciation. Consider these words with word-final <mb> spelling [m]: bomb, climb, comb, dumb, lamb, succumb, tomb, womb. Rather than calling all those final <b>'s silent letters that mark some ghostly unit of silence, in this analysis they are treated as simplifications – that is, cases where earlier pronunciations simplified but the spelling has remained the same. Thus, in these words we say that [m] is spelled <mb>, in cupboard [b] is spelled <pb>, and in subpoena [p] is spelled <bp>. One holdout: the letter <h> does occur as a true silent letter in words like honest, honor, hour and after <x>, as in exhaust, exhibit, exhort. See <h> and [h]

This analysis does, however, recognize silent vowel letters – most noticeably, silent final <e> as in time where it serves the diacritic function of marking a long vowel, or clothe where it marks a voiced <th>, ounce where it marks a soft <c>, or bronze, clause, league, active, where it insulates a letter that normally doesn’t occur at the end of word-final bases. On the other hand, some silent <e>'s serve no diacritic function at all, as in fixed and ashamed. See <gh>. For more on <gh> in words like weigh and neighbor see “A Note on Silent Letters” from Spelling for Learning.

**Simple Addition.** The Rule of Simple Addition is the single most important and general rule in spelling. It states that unless you know of some reason to make a change – such as twinning or <e> deletion – elements simply add together, or concatenate, with no changes – as one of my students said years ago, like pop-beads. Far and away most of the times elements join through simple addition – as in unearthly and repainted.

**Simplifications.** Simplifications retain the original longer spellings of one-time blends that have simplified over time to single consonant sounds:

<table>
<thead>
<tr>
<th>Blend</th>
<th>Original Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;cht&gt;</td>
<td>[t] as in yacht</td>
</tr>
<tr>
<td>&lt;ft&gt;</td>
<td>[f] as in often</td>
</tr>
<tr>
<td>&lt;ght&gt;</td>
<td>[t] as in light</td>
</tr>
<tr>
<td>&lt;gn&gt;</td>
<td>[n] as in sign</td>
</tr>
<tr>
<td>&lt;kn&gt;</td>
<td>[n] as in knight</td>
</tr>
<tr>
<td>&lt;ld&gt;</td>
<td>[d] as in could</td>
</tr>
</tbody>
</table>
Some analyses would treat the extra letters in simplifications as silent letters or fossils.

**Simplifications with lost [t], [l], and [p].**

**Lost [t]:** Usually <ft> spells [ft] as in *after, craft, fifteen, fifty, gift, left, lofty, oft, swift*. But it spells [f] with a lost [t] in *often* – though there is a less common variant pronunciation with [ft]. These variant pronunciations have been around since the 16th century, though [ft] seems to be gaining ground, probably due to spelling pronunciation. The [t] is also lost in other simplifications, especially before the syllabic consonants [n] and [l]: *fasten, listen, soften; castle, epistle, thistle, wrestle*, and in *chestnut, Christmas, mortgage*.

**Lost [l]:** Usually <ld> spells [ld]: *bewilder, children, field, held, old, soldier, world*. But in *could, should, would* and their derivatives *couldn't, shouldn't, wouldn't*, <ld> spells [d] with a lost [l]. See Analogy. There is also a lost [l] in *solder*. Usually <lf> spells [lf]: *fulfill, golf, herself, olfactory, shelf, sulfur, twelfth, werewolf*. But the [l] is lost in *behalf, calf, half*. Usually <lk> spells [lk]: *alkaline, bulk, folk, polka, silken*, but the [l] is lost in *chalk, Norfolk, sidewalk, stalk, talk, walk*. The only known case where [l] is lost in an <ln> spelling is *Lincoln*.

**Lost [p]:** Usually <ps> spells [ps]: *apocalypse, collapse, eclipse, epilepsy, keeps, perhaps, upset*. But in several words, all from Greek spelled with the Greek letter psi, <ψ>, originally pronounced [ps], the [p] is lost in English: *parapsychology, psalm, psalter, pseudoscience, psyche, psychiatry, psychic, psychology, psychotic, psychosomatic, psychotherapy*. The [p] is also lost in *receipt* due to a 14th century Latinization of the Middle English *receite*, from the Latin *recepta*. And the [p] is lost in a number of [pt] simplifications of Greek adaptions, mostly from the scientific-technical register: *pterodactyl, ptisan, ptomaine, ptosis, ptyalin*. A [p] is also lost in the respelling of the Gaelic *tarmachan* to *ptarmigan*, apparently due to the mistaken notion that it was a variation of the Greek base *pter* “wing” as in *pterodactyl* and *helicopter*. There is a lost [p] in the simplification <pb> in *cupboard, clapboard, raspberry, Campbell*. This
loss is due to the fact that when the the two bilabial stops [p] and [b] concatenate, the first tends to be lost. Compare this with loss of [b] in subpoena. (In the concatenation <mbp> in bombproof and thumbprint the lost [b] is due to <mb> simplification in bomb and thumb.)

Soft and hard <c> and <g>. The letter <c> spells the sound [s] when it is followed by the letters <e, i, y>: chance, chancing, chancy. A <c> that spells [s] is called soft <c>. The letter <c> spells the sound [k] before any other letter and at the end of a word: career, discuss, nucleus, critic, arc. A <c> that spells [k] is called hard <c>. The letter <g> usually spells the sound [j] when it is followed by the letters <e, i, y>: range, ranging, rangy, and is called soft <g>. The letter <g> spells the sound [g] before any other letter and at the end of a word. But it also spells [g] before <e, i, y> in some native words: get, girl, bogy – and in one pronunciation of the Greek base gyn “woman,” as in gynecology. When <g> spells [g], it is called hard <g>.

Spelling Bees. Spelling bees are named after the honey bee in reference to the bee’s social nature. Originally this word bee was used to refer to work parties, especially among farm families, where neighbors joined together in some project – as in quilting-bees, (barn-)raising-bees, husking-bees. Later it came to mean a gathering for a competition among spellers. For more information about becoming involved with spelling bees, you can go to www.spellingbee.com.

Spelling Demons. There are many different lists of spelling demons, but very little agreement among them. In his Spelling: An Element in Written Expression R. L. Hillerich reports that in a 398-word composite list from four different lists of spelling demons, more than 70% appeared in only one of the four lists (Merrill, 1976, p. 64). An even larger composite demons list appears in the Other Problem Spellings field of the CommonWords data table, where 639 words that appear in at least one list of demons are marked with an exclamation point. See also Misspellings.

Spelling pronunciation. Spelling pronunciation is a change in pronunciation that brings it more in line with the spelling. It often results in the introduction of consonant sounds that were missing in the traditional pronunciation – for example, often with [t], clothes with [th], salmon and palm with [l], arctic with two [k]'s, ski with [sk] rather than the original [sh]. Sometimes it results in the re-expansion of pronunciations that had traditionally been simplified: waistcoat as [wâst'kôt] rather than the older [wēs'kôt], cupboard with a [p], forehead as [fôr'heid] rather than the older [fôr'îd].

Standard Written English. After the reestablishment of English as the language of the land in the 14th and 15th centuries, a written standard developed in the Early Modern English period. The standard was based on the London dialect and was helped along by the invention of the printing
press and the work of the Chancery office in London, which produced
government documents that were circulated nationwide. Standard Written
English went far towards standardizing English spelling, but since the
spoken language changes faster than the written, we are left with a spelling
system that at times does a more efficient job of spelling Early Modern
English than the English of today.

**Stems.** A stem is any element or string of elements to which we are going
to add or from which we are going to subtract any elements. Every stem
must contain at least one base, but it can have any number of prefixes or
suffixes, including none at all. Stems, like bases, can be either bound or
free. So all of the following combinations can be stems – in this case, free
stems:

<table>
<thead>
<tr>
<th>Combinations</th>
<th>Example Stems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free Base</td>
<td>paint</td>
</tr>
<tr>
<td>Prefix + Free Base</td>
<td>repaint</td>
</tr>
<tr>
<td>Free Base + Suffix</td>
<td>painted</td>
</tr>
<tr>
<td>Prefix + Free Base + Suffix</td>
<td>repainted</td>
</tr>
</tbody>
</table>

And we can go on adding elements, especially prefixes and suffixes, as in
free stems like *unrepainted*.

Stems can also contain bound bases, as in the following examples,
with the bound base *spect*:

<table>
<thead>
<tr>
<th>Combinations</th>
<th>Example Stems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefix + Bound Base</td>
<td>respect</td>
</tr>
<tr>
<td>Prefix + Bound Base + Suffix</td>
<td>respectful</td>
</tr>
<tr>
<td>Prefix + Prefix + Bound Base + Suffix</td>
<td>disrespectful</td>
</tr>
<tr>
<td>Prefix + Prefix + Bound Base + Suffix + Suffix</td>
<td>disrespectfully</td>
</tr>
</tbody>
</table>

Stems can also be bound, as in stems like *spect*, *spectful* and
*spectfully*, none of which is a word. Sometimes even a stem with a free
base can be bound. For instance, In the word *unblinkingly*, *unblinking* is a
free stem because it's also a word, but *unblink* is bound, since we do not
have the word *unblink*.

**Stress.** Stress (sometimes called accent) refers to what is usually called
the loudness of a certain vowel sound – though it is often described as
intensity or prominence or force. Some linguists distinguish four degrees of
stress: primary, secondary, tertiary, and weak, but usually only three
degrees are distinguished – primary, secondary, and weak – as in the
pronunciation respellings given in most dictionaries. In the analysis of
syllable structure in CommonWords, only two degrees are noted, stressed
and unstressed. Other areas of CommonWords recognize three degrees – primary marked with a following high vertical, secondary with a following low vertical, and weak with no mark: [flā'mīng], [bīg,hēdīd], [ə-būv].

**Strings.** In CommonWords a string is a sequence of vowel and consonant letters, the first vowel of which is usually spelling a stressed vowel sound. Four different strings are recognized:

- **VCV vs. VCCV strings.** A VCV string consists of a single vowel letter spelling a stressed long vowel sound followed by a single consonant letter, which is in turn followed by a vowel letter – often a silent final <e>: rated, ceded, writing, hope, dune. A VCCV string consists of a single vowel letter spelling a stressed short vowel sound followed by two consonant letters – often a doublet – as in gander, lettuce, sadden, pumpkin. Thus we have contrasts like later vs. latter, biter vs. bitter, coma vs. comma, dining vs. dinning, fiber vs. fibber. See [VCV vs VCCV Contrasts](#).

- **VCCle vs. VCle strings.** A VCCle string consists of a single vowel letter spelling a stressed short vowel sound followed by two consonant letters – often a doublet – followed in turn by a word-final <le>. Some examples: battle, gamble, pebble, gentle, thimble, topple, supple. A VCle string consists of a single vowel letter spelling a stressed long vowel sound followed by a single consonant letter followed <le> – as in ladle, noble, ogle, quadruple. Thus we have the following contrasts: riffle vs. rifle; ruble vs. rubble.

- **CVC# strings.** A CVC# string consists of a word-final consonant letter preceded by a single vowel letter, with the vowel letter spelling a short vowel – as in homonym, lunatic, thermostat.

- **V.V strings.** A V.V string consists of two vowel letters spelling two vowel sounds separated by a syllable boundary. The first vowel in a V.V string will be long even if it is unstressed, as in create [krē-ət] or anemia [ə-nē-mē-a], as well as in words like lion and giant where the first vowel is stressed. Notice that the first vowel in a V.V string is by definition in an open syllable.

**Suffix -ic)1 Rule.** See [Suffix Rules](#).

**Suffix -ity) Rule.** See [Suffix Rules](#).

**Suffixes.** There are two kinds of suffixes: inflectional and derivational. In dictionaries inflected forms are usually listed towards the beginning of an entry, derived forms towards the end.

- **Inflectional suffixes.** Inflectional suffixes add meanings to the stem that help answer questions like "One or more than one?" "Whose?" "How
much?" and "When?" For instance, the inflectional suffix -ed)1 answers the question "When?", and the suffix -s)3 answers the question "One or more than one?" In the sentence "George painted three of the chairs," -ed)1 adds the meaning "in the past" to the verb paint; -s)3 adds the meaning "more than one" to the singular noun chair. Thus, -ed)1 is a past tense verb inflectional suffix; -s)3 is a noun plural inflectional suffix.

Among regular verbs there are two common 3rd person singular present tense suffixes: -s)2 and -es)1. We add -es)1 to verbs that end with one of four sounds: [s], [z], [ch], or [sh], and after the <y> to <i> replacement – as in kisses, buzzes, watches, wishes, tries. Fixes takes -es)1 because the <x> spells the combination sound [ks]. These are the inflectional suffixes for regular verbs:

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Inflectional Sense</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>-s)2, -es)1</td>
<td>3rd person singular present tense “action undertaken”</td>
<td>“She calls.”</td>
</tr>
<tr>
<td>-ed)1</td>
<td>Past tense “in the past”</td>
<td>“She called yesterday.”</td>
</tr>
<tr>
<td>-ed)1</td>
<td>Past participle “action completed”</td>
<td>“She has called often.”</td>
</tr>
<tr>
<td>-ing)1</td>
<td>Present participle “action continuing”</td>
<td>“She is calling now.”</td>
</tr>
</tbody>
</table>

Notice that what we are calling present tense is actually a kind of recurrent or routine tense: “She calls” could mean that she may be calling right now in the present, but more often it means that she has called in the past, and can be expected to call in the future.

Nonregular, so-called strong verbs from Germanic show past tense and past participle inflection through changes in their vowel. See Ablaut. But they show the 3rd person present singular with -s)2 and the present participle with -ing)1: sing, sang, sung, sings, singing; swim, swam, swum, swims, swimming.

There are also two common noun plural suffixes: -s)3 and -es)2. See Noun plurals.

The inflectional suffixes -er)02 and -est)1 form comparatives and superlatives, adding the senses “more” and “most”: dark, darker, darkest; slow, slower, slowest.

Derivational Suffixes. Inflectional suffixes add layers of meaning to words. But derivational suffixes usually change the entire function of the words to which they are affixed – that is, change their part of speech, adding information like "This is an adjective" or "This is an adverb." A derivational suffix is basically any suffix that is not an inflection – and there are a lot of them. Though the situation is more complicated, the following
illustrates some of the uses of several derivational suffixes:

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Change</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>-able</td>
<td>verbs into adjectives</td>
<td>agree, agreeable</td>
</tr>
<tr>
<td>-ance</td>
<td>verbs into nouns</td>
<td>guide, guidance</td>
</tr>
<tr>
<td>-ancy</td>
<td>adjectives or verbs into nouns</td>
<td>buoy, buoyancy</td>
</tr>
<tr>
<td>-al1</td>
<td>nouns into adjectives</td>
<td>classic, classical</td>
</tr>
<tr>
<td>-al2</td>
<td>verbs into nouns</td>
<td>propose, proposal</td>
</tr>
<tr>
<td>-ful</td>
<td>nouns or verbs into adjectives</td>
<td>delight, delightful</td>
</tr>
<tr>
<td>-ion1</td>
<td>verbs into nouns</td>
<td>act, action</td>
</tr>
<tr>
<td>-ity1</td>
<td>adjectives into nouns</td>
<td>extreme, extremity</td>
</tr>
<tr>
<td>-ize</td>
<td>nouns into verbs</td>
<td>critic, criticize</td>
</tr>
<tr>
<td>-less2</td>
<td>nouns into negative adjectives</td>
<td>hope, hopeless</td>
</tr>
<tr>
<td>-ly1</td>
<td>adjectives into adverbs</td>
<td>close, closely</td>
</tr>
<tr>
<td>-ly3</td>
<td>nouns into adjectives</td>
<td>friend, friendly</td>
</tr>
<tr>
<td>-ment1</td>
<td>verbs into nouns</td>
<td>achieve, achievement</td>
</tr>
<tr>
<td>-ness1</td>
<td>adjectives into nouns</td>
<td>kind, kindness</td>
</tr>
<tr>
<td>-th1</td>
<td>cardinal numbers into ordinals</td>
<td>eleven, eleventh</td>
</tr>
<tr>
<td>-th2</td>
<td>verbs or adjectives into nouns</td>
<td>grow, growth; warm, warmth</td>
</tr>
</tbody>
</table>

Some derivational suffixes do not change an adjective or noun’s part of speech but rather turn it into a related noun:

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Sense</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ess1</td>
<td>“Feminine”</td>
<td>god, goddess</td>
</tr>
<tr>
<td>-hood</td>
<td>“Condition, instance of a condition”</td>
<td>child, childhood</td>
</tr>
<tr>
<td>-ism</td>
<td>“Action, characteristic quality”</td>
<td>terror, terrorism</td>
</tr>
<tr>
<td>-ist1</td>
<td>“One connected with a thing”</td>
<td>violin, violinist</td>
</tr>
<tr>
<td>-ster1</td>
<td>“One that is associated with”</td>
<td>gang, gangster</td>
</tr>
</tbody>
</table>

**Suffix Rules.** The two main suffix rules concern the suffixes -ity1 and -ic1. The **suffix -ity rule** states that the suffix -ity1 is regularly preceded by a vowel that is stressed and short, even if it heads a VCV string. The -ity1 rule motivates shortening of the stem vowel, as in sane with [ā] vs. sanity with [ã], and it also motivates a shift of stress onto the vowel immediately preceding it, as in civil vs. civility, [sɪˈvɪl] vs. [sɪˈvɪlɪ-tē]. Other examples: active, activity; mental, mentality; cave, cavity; obscene, obscenity; electric, electricity; profane, profanity; extreme, extremity; public, publicity.

The **suffix -ic1 rule** states that -ic11 is regularly preceded by a vowel that is stressed and short, even if the vowel is unstressed or long in the stem: athlete, athletic; state, static; atom, atomic; parasite, parasitic; demon, demonic; patriot, patriotic.

**Suppose, supposed; use, used.** Suppose is a verb; supposed is the past participle of that verb used as an adjective meaning “intended as true, required.” The major spelling problem is when the verb suppose is used in phrases that call for the adjective supposed, as in the mistaken *“He was
suppose to go.” A similar problem exists with the verb use and its past participle used, as in the mistaken **“He use to go on Saturdays”** instead of “He used to go on Saturdays.”

**Surnames.** One’s surname – or last name – is the name of one’s family, in contrast with one’s first, or given, name, which identifies an individual in that family. Surnames can have many different sources and original meanings; often those given below are just thought to be likely.

Many surnames originally showed descent, in English nearly always from a father – for instance, Jacobson “son of Jacob” or Petersen “son of Peter.” Such surnames are called *patronymics* “fathers’ names.” Besides <son> and <sen> other additions that mean “of” or “son of” are

- the possessives <s> and <es> as in Roberts and Hughes;
- <mac>, often contracted to <mc>, as in McDonald and McOwen;
- <ez> as in Hernandez and Gomez;
- <o'> as in O’Brien and O’Connor, which is lost in Sullivan;
- <fitz> as in Fitzgerald and Fitzpatrick;
- <ap> as in Price, contracted from the Welsh Ap Rhys;
- <ben> as in Benjamin.

Among the 100 most common surnames in the 2000 U.S. census, the largest group are patronymics. The following list is in descending order:

<table>
<thead>
<tr>
<th>Johnson</th>
<th>Robinson</th>
<th>Kelly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Williams</td>
<td>Perez</td>
<td>Diaz</td>
</tr>
<tr>
<td>Jones</td>
<td>Sanchez</td>
<td>Richardson</td>
</tr>
<tr>
<td>Davis</td>
<td>Adams</td>
<td>Watson</td>
</tr>
<tr>
<td>Garcia</td>
<td>Nelson</td>
<td>James</td>
</tr>
<tr>
<td>Rodriguez</td>
<td>Ramirez</td>
<td>Reyes</td>
</tr>
<tr>
<td>Wilson</td>
<td>Roberts</td>
<td>Brooks</td>
</tr>
<tr>
<td>Martinez</td>
<td>Phillips</td>
<td>Hughes</td>
</tr>
<tr>
<td>Anderson</td>
<td>Evans</td>
<td>Price</td>
</tr>
<tr>
<td>Thomas</td>
<td>Torres</td>
<td>Myers</td>
</tr>
<tr>
<td>Hernandez</td>
<td>Collins</td>
<td>Sanders</td>
</tr>
<tr>
<td>Jackson</td>
<td>Edwards</td>
<td>Morales</td>
</tr>
<tr>
<td>Thompson</td>
<td>Flores</td>
<td>Powell, from Ap Hywel</td>
</tr>
<tr>
<td>Lopez</td>
<td>Morris</td>
<td>Sullivan</td>
</tr>
<tr>
<td>Gonzalez</td>
<td>Rogers</td>
<td>Ortiz</td>
</tr>
<tr>
<td>Harris</td>
<td>Peterson</td>
<td>Jenkins</td>
</tr>
<tr>
<td>Lewis</td>
<td>Gomez</td>
<td>Gutierrez</td>
</tr>
</tbody>
</table>

Many surnames, especially those ending with <er> or <or>, indicate occupations for which the original bearer of the name was known – such as Baxter (feminine of Baker), Carpenter, Sawyer, Barber. Among the most common in North America, again in descending order:

Some surnames indicate qualities or characteristics of the original bearer, like a nickname, as in Armstrong, Becket(t) “little beak”, Black, Blanchard “whitish.” Other common ones:

**White**  **Cox “youth, cocky”**  **Howard “brave”**
**Allen “little rock”**  **Bell “handsome”**  **Gray**
**Young**  **Long**  **Russell “red hair”**  **Mitchell “like God”**
**Green**  **Campbell “crooked mouth”**  **Morgan “sea bright”**  **Reed “red faced or red haired”**

Some surnames indicate the locality with which the original bearers were identified, as in these common ones:

**Moore “moor”**  **Rivera**  **Lee “lea, meadow”**  **Wood**  **Scott “Scotland”**  **Ross “upland, moorland”**
**Hall**  **Perry “pear tree”**  **Hill**  **Barnes**

Other common surnames are **Martin “of Mars, god of war”; the Vietnamese Nguyen from a Chinese surname, forced on many Vietnamese people by rulers’ fiats; Murphy “descendant of a sea warrior”; Bennett “after St. Benedict”; Cruz “cross.”** And, of course, the meanings and sources of many surnames remain a mystery.

For more see OCEL, “Patronymics” and “Personal Names.” A short, but good, dictionary is Basil Cottle’s paperback The Penguin Dictionary of Surnames (1967, 1984).

**Syllabic consonants.** A *syllable* always contains one and only one peak of sound. In the huge majority of cases those peaks are vowel sounds spelled with vowel letters, but occasionally the *nasals and liquids* [m], [n], [l], and [r] can provide the peak without a vowel letter, usually at the end of free stems. These consonant sounds are called *syllabic consonants*. For instance, the word *chasm* has two syllables but only one vowel letter.
When the letter \( <m> \) follows an \( <s> \) and comes at the end of a word, it spells a syllabic consonant that functions as if it were a vowel. Other instances are in *algorithm*, *organism* – and all the other words with the suffix *-ism*.

Other letters that can spell syllabic consonants are \(<n>\), \(<l>\), and \(<r>\). Notice that in *button*, *brittle*, and *butter*, you do not hear any vowel sounds in the final syllables, just the consonant sounds [n], [l], and [r].

**Syllables.** Syllables are units of the spoken language. That is one of the ways they differ from *elements*, which are units of the written language. A syllable is a section of a spoken word that contains one and one only vowel sound – or in a few cases one and only one syllabic consonant. Syllables can be stressed or unstressed, and open or closed. Open syllables end with a *vowel sound*; closed syllables end with a *consonant sound*. Closed syllables usually contain short vowel sounds, as in *bat* and *monk*, though closed syllables at the end of words can contain long vowels, especially those that end with *silent final* \(<e>\), as in *hope* and *engage*. Open syllables usually contain a long vowel sound, as in *agent* and *sky*. Reduced vowels can occur in both open and closed syllables, like the schwas in *above* and *item*.

In dictionaries the syllabification of the entry word is given in the pronunciation respelling, usually in parentheses right after the word. But dictionaries also divide the entry word itself into units that fall somewhere between regular syllables and elements – think of them as written syllables. They are provided to show where the word can be divided and hyphenated at line’s end. These divisions very often do not agree with the actual syllabification of the word or with its *explication* into elements. See “A Note on Syllables” from *Spelling for Learning*.

**Synonyms, antonyms, pseudonyms, heteronyms, acronyms.**

- **Synonyms** are two or more words that have the same or nearly the same *content* – such as *lawyer* and *attorney*, or *throng* and *crowd*. Though their contents are close, most synonyms actually produce slightly different shades of meaning, so they must be used with care.
- **Antonyms** are words that have opposite or nearly opposite contents – like *hot* and *cold*, or *happy* and *sad*.
- **Pseudonyms** are fictitious names. A pseudonym used by a writer is called a *pen name* or *nom de plume*, such as Samuel Clemens’ pen name *Mark Twain*.
- **Heteronyms** have the same spellings but different contents and pronunciations – such as *bass*1 “a low singing voice” pronounced [bəs] and *bass*2 “a fish” pronounced [bæs]; or *tinged*1 “made a tinging noise” pronounced [tɪŋd] and *tinged*2 “slightly colored” pronounced [tɪŋd]. *Heteronym* is another name for *homograph*.
- **Acronyms** are formed from the initials of a group of words. Some acronyms are pronounced like regular words – for instance, *NATO* and *UNICEF*, but others are spelled out – *UN* and *ROTC*, though on college
Third Syllable (or Third Vowel) Rule. The Third Vowel Rule says that the third (or fourth) vowel sound from the end of the word will be short if it is stressed. Most instances of the Third Vowel Rule are words adapted from Latin, and they reflect the way Latin pronunciation was taught in British schools during the late Middle Ages and Renaissance. A few instances, like holiday from “holy day,” are native English words, which reflect the fact that in Old English there was a strong tendency to shorten long vowel sounds in syllables three or more places from the end of a word. Another native English example is midwifery with [wɪf] from midwife with [wɪf].

In the following pairs, the first, the stem or a closely related shorter word, has a long vowel, while the second is a longer derived or related word that illustrates the Third Vowel Rule. In these words we are treating -ion as having two syllables, as it did back when the Third Vowel Rule affected it:

<table>
<thead>
<tr>
<th>Shorter Words</th>
<th>Longer Words</th>
<th>Shorter Words</th>
<th>Longer Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>compete</td>
<td>competitor</td>
<td>legal</td>
<td>legacy</td>
</tr>
<tr>
<td>crime</td>
<td>criminal</td>
<td>nation</td>
<td>national</td>
</tr>
<tr>
<td>decide</td>
<td>decision</td>
<td>nature</td>
<td>natural</td>
</tr>
<tr>
<td>sane</td>
<td>sanity</td>
<td>navy</td>
<td>navigate</td>
</tr>
<tr>
<td>explain</td>
<td>explanatory</td>
<td>pose</td>
<td>positive</td>
</tr>
</tbody>
</table>

The Third Vowel Rule holds in hundreds and hundreds of words, perhaps thousands, but there are two familiar local rules that can preempt it:

• The long vowels [ʊ] and [yʊ], as in cuticle, enthusiast, fumigate, jubilee, luminous, mutilate, punitive, and unicorn; and

• when the head vowel of a V.V string falls three (or four) places back in a word, it will resist shortening by the Third Vowel Rule: dialect, iodine, piety, psychiatrist, reliable, violate, hyacinth, kaolin, peony.

Through, thorough. Both through and thorough come from the Old English word þurh “through.” By late Old English þurh had developed the two-syllable expansion þuruh with the extended meaning “through, end to end,” which became our adjective thorough “complete.” This same expansion from one to two syllables occurred in several words – including our words borough, furrow, borrow, sorrow. For <þ> see Runes.

Triplets. In English spelling there is a strong constraint against triplets – that is, strings of three identical letters. Thus, shrill+ly) does not equal *shrrllly and wee+est) does not equal *weeeest. I am not aware of any holdouts to this constraint. We can say that shrilly explicates to either...
shrill+y)4 or to shrill+ly)1, the suffix -y)4 being a contraction of -ly)1 used after [l], as in notably, visibly, etc. By the same token, weest explicates either to wee+est)1 or to wee+st)5, -st)5 being a contraction of -est)1 as in first, most, etc.

There is a weak constraint against doublets within larger clusters. Thus, offal, which etymologically is off + fall has a deleted <f> (and <l>) off+fall. This constraint also explains the deletions in ascend, which equals (ad+s+scend to avoid *asscend, with the <ss> doublet in the larger <ssc> cluster. Similarly we have aspect (ad+s+spect, distant (dis+stant, and distinct (dis+stinct and several others. However, there are several holdouts to this weak constraint – for instance, in some words that end with the suffixes -ful) or -ness): blissful, illness, successful, wellness, and several others.

**Twinning.** You twin the final consonant of a stem if you are adding a suffix that starts with a vowel, and if the stem ends in a single vowel sound and letter followed by a single consonant sound and letter, and if there is at least secondary stress on the final vowel sound of the stem both before and after you add the suffix. For example:

\[
\text{begin } + \text{er})1 = (\text{be}+\text{gin}3+n+\text{er})01 = \text{beginner} \\
\text{commit } + \text{ing}) = (\text{com}+\text{mit}1+t+\text{ing})1 = \text{committing}
\]

But dead + er)02 = deader with no twinning because there are two vowel letters preceding the final consonant. Rich + er)02 = richer with no twinning because the stem word ends with two consonant letters. And limit + ed) = limited, symbol+ic)1 = symbolic, and inference) = inference – all with no twinning because there is not stress on the final syllable of the stem word both before and after the suffix is added. For more details see “Twinning” from Spelling for Learning.

**Use, used.** See Suppose, supposed.

**Variant Spellings.** In his Variant Spellings in Modern American Dictionaries (1973) Donald W. Emery lists about 2,000 words that have variant spellings. Three points to remember about variant spellings are (i) all are equally correct, (ii) not all are equally common, and (iii) not all are equally regular. In general, when correcting tests and papers, I recommend a generous approach: If a spelling is listed as a variant in at least one dictionary, it’s correct – though it’s probably a good idea to point out to students that uncommon variants will often be treated as incorrect by readers. Further tempering that generosity, if one of the variants is more regular than the other or others, I recommend preferring the more regular one. By regular here I mean simply that the spelling better agrees with the patterns and rules of English spelling. By preferring regularity we are gently encouraging the evolution of our spelling system to more regularity and ruliness. See Principle of Preferred Regularity.
**VCV vs. VCCV Contrasts.** VCV vs. VCCV contrasts occur in several common pairs of words and can give spellers problems:

- biter, bitter
- coma, comma
- diner, dinner
- filing, filling
- fury, furry
- later, latter
- liter, litter
- muter, mutter
- planing, planning
- raged, ragged
- supper, super

VCV vs. VCCV contrasts also occur in less common pairs of words:

- coper “one who copes”, copper
- fluter “one who plays the flute”, flutter
- mater “Latinate ‘mother’”, matter
- puling “whining”, pulling
- shining, shinning “shinnying”
- wining as in “wining and dining”, winning

Similar single-consonant vs. double-consonant contrasts occur, but due to shifts in stress they are not regular VCV vs VCCV contrasts:

- below, bellow
- cares, caress
- desert, dessert
- discus, discuss
- posses (pl. of posse), possess
- princes, princess

There are a very few single-consonant vs. double-consonant contrasts with no change in the pronunciation of the vowels: canon, cannon; caries “tooth decay”; carries; and Mary, marry.

**Velar sounds.** Velars are consonant sounds articulated with the back of the tongue at or near the soft palate, or velum, at the rear of the mouth.

**Verbs: regular and irregular.** Almost all verbs change their form to indicate action in the past. Most verbs are regular, which means that they form their past tense and past participle by adding the suffix -ed):1 call (pres.), called (past), called (past part.); buzz, buzzed, buzzed; look, looked, looked; tow, towed, towed; pass, passed, passed. But there are more than 250 irregular verbs – that is, verbs that do not form their past tense and/or their past participles with -ed):1. The following are some of the different types of irregular verbs: A few, like cut, cost, hit, hurt, shut, do not make any changes in their past tense or past participle. Several, the so-called strong verbs, form their past tense by changing the vowel sound and spelling and their past participle by sometimes changing the vowel and always adding the suffix -en):2 or -n):1: freeze, froze, frozen; speak, spoke, spoken; draw, drew, drawn; fall, fell, fallen, eat, ate, eaten; rise, rose, risen. See Ablaut. Others mark the past tense and past participle in the same way, but with pronounced changes in spelling and pronunciation from the infinitive or basic form: bring, brought, brought; buy, bought, bought; catch,
caught, caught; teach, taught, taught; think, thought, thought.

Regular and irregular verbs all form the 3rd person singular present tense and the present participle with -s)2 or -es)1 and -ing)1, as in thinks, goes; thinking, going. See Suffixes.

Vikings. The Vikings were seafarers from Norway, Sweden, and Denmark who from the late 8th century to the Norman Conquest in the 11th raided northern Europe and England. Those who settled in northern France in the 10th century became the Normans. There were earlier landings and raids, but the Vikings’ first major raid on England was in 793, when they sacked and destroyed the Christian abbey on the island of Lindisfarne off England’s northeast coast. By the mid 9th century they had overrun and begun settling eastern and northern England. By the late 9th century and after repeated battles Alfred the Great, king of Wessex, defeated the Vikings and brought peace to Wessex in southern England and to Mercia in the central south and west. In the late 10th century the Viking attacks began again and continued through the early 11th century, until the Norse had settled in and held control of much of eastern England. In the early 11th century there were four Norse kings of England: Sweyn Forkbeard, Canute the Great, Harold Harefoot, and Harthacnut.

The interaction between Old English and Norse and Anglo-Norse had significant influence on the evolution of English.

Voiced vs. voiceless (unvoiced) sounds. Sixteen consonants come in pairs in which the two sounds are identical except that one is voiced and the other is voiceless. A voiced sound is pronounced with the vocal cords vibrating, a voiceless sound is pronounced with no vibration:

<table>
<thead>
<tr>
<th>Voiceless</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>[p] as in pop</td>
<td>[b] as in bob</td>
</tr>
<tr>
<td>[t] as in tat</td>
<td>[d] as in dad</td>
</tr>
<tr>
<td>[k] as in kick</td>
<td>[g] as in gig</td>
</tr>
<tr>
<td>[f] as in fluff</td>
<td>[v] as in verve</td>
</tr>
<tr>
<td>[s] as in scarce</td>
<td>[z] as in czars</td>
</tr>
<tr>
<td>[th] as in thin</td>
<td>[th] as in then</td>
</tr>
<tr>
<td>[sh] as in shush</td>
<td>[zh] as in azure</td>
</tr>
<tr>
<td>[ch] as in church</td>
<td>[j] as in judge</td>
</tr>
</tbody>
</table>

Nasals, liquids, semivowels, and vowels are all voiced.

Sometimes, even if you have trouble hearing the difference between, say, voiceless [th] and voiced [th], as in thin and then, you can feel the
difference: Put your fingers lightly on your throat just up under your chin and say thin. You should feel nothing as you pronounce the [th]. Then say then. You should feel some vibration in your throat as you pronounce the voiced [th]. The vibrations are caused by your vocal cords.

**Vowel lengthening and shortening.** In earlier English the distinction between long and short vowels was quite different from what it is today. It was a distinction in sound quantity, or duration, as the words long and short suggest: Long <a> was the same vowel sound as short <a>, just held longer. Thus, in earlier English short <a> was an “ah”-like sound; long <a> was more like “aah.” Today long and short refer not so much to quantity, or duration, but to the quality of the vowel sound, so today's short <a>, [ã] as in fat, is different in sound quality from long <a>, [a] as in fate, with no consideration of quantity. Today long vs. short is not a phonetic but an orthographic distinction.

In Old English long vowels could occur in either open or closed syllables, as could short vowels. The evolution of Old English into Middle English was marked by the systematic lengthening of short vowels in open syllables while long vowels in closed syllables were systematically shortened. The following are words with long vowels that were short in Old English open syllables: bacon, eat, hope, make, meat, sake, smoke, steal. Words with short vowels that were long in Old English closed syllables include ask, kept, hid, soft.

The lengthening of Old English vowels in open syllables led to the convention of using silent final <e> as a long vowel marker. This final <e> was all that was left of a number of Old English inflectional suffixes and was originally pronounced [e], usually putting the preceding vowel in an open syllable, thus tending to lengthen it. After the final <e> fell silent in late Middle English, it was still there in the spelling to complete the VCV string.

There were some other more localized and less systematic lengthenings: Some short Old English vowels lengthened before <nd>, <ld>, and <mb>: bind, blind, find, grind, hind, kind, mind, rind, wind (vb.); child, wild, mild; climb, tomb. For more on vowel lengthening and shortening and the use of silent final <e>, see AES, pp. 97-99, “A Short History of the VCC/VCV Contrast” and “The Development of Silent Final e.”

**Vowel letters.** See Consonant and vowel letters.

**Vowel sounds.** English vowels are usually displayed in a chart whose rows indicate the tongue’s position relative to the roof of the mouth (high, mid, low) and whose columns indicate the part of the tongue that is raised (front, central, back). The simple vowels recognized in our analysis are as follows:
<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td>[i] beet /i/</td>
<td></td>
<td>[u] boot /u/</td>
</tr>
<tr>
<td></td>
<td>[i] bit /i/</td>
<td></td>
<td>[u] book /u/</td>
</tr>
<tr>
<td><strong>Mid</strong></td>
<td>[ä] bait /e/</td>
<td>[ü] but /ü/</td>
<td>[ö] boat /o/</td>
</tr>
<tr>
<td></td>
<td>[ê] bet /e/</td>
<td>[œ] alone /œ/</td>
<td></td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>[ä] bat /æ/</td>
<td></td>
<td>[ä] bother /a/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[ö] cot /o/</td>
<td>[ö] caught /o/</td>
</tr>
</tbody>
</table>

Characters in diagonals are equivalent IPA symbols. Notice that [ü] is close to [û], and that [ö] and [ä] are close to [ô].

Ten vowel sounds come in five short/long pairs:

<table>
<thead>
<tr>
<th>Short</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ä] bat /æ/</td>
<td>[ä] bait /eɪ/</td>
</tr>
<tr>
<td>[ê] bet /e/</td>
<td>[ê] beet /i/</td>
</tr>
<tr>
<td>[i] bit /i/</td>
<td>[i] bite /aɪ/</td>
</tr>
<tr>
<td>[ö] got /o/</td>
<td>[ö] goat /oʊ/</td>
</tr>
<tr>
<td>[ü] but /u/</td>
<td>[ü] boot /u/</td>
</tr>
</tbody>
</table>

There is a second long <u> recognized in our analysis, [yû], /ju/ in IPA, as in cute, contrasting with [û] in coot. It is essentially a diphthong. See Vowel lengthening and shortening. (EDLL, OCEL, AES, 201-06, 210-12)

**Wear, ware, where; were, we’re.** In this group wear, ware, where are homophones, which can always be problems for spellers; were and we’re are near homographs, which shouldn’t be too problematic – if students remember that apostrophe in the contraction we’re.

These homophones have many historical sources and are often the product of changes in vowel pronunciations and spellings in earlier English, which are complicated further by the effects of [r] on preceding vowels. Alas, I know of little specific historical information that would be useful in the elementary or secondary classroom. But on the assumption that it is always better to know your enemies, here are the most common instances of the homophonic spellings <ear>, <are>, and <ere>:

Like wear: bear, pear, swear, tear “rip” (vb.).
Like ware: bare, care, compare, dare, declare, fare, flare, glare, hare, mare, prepare, rare, scare, share, snare, spare, square, stare.
Like where: ere, plus there and where and their compounds, such as thereafter, wherever, and the like.

Webster, Noah. In 1783 the American Noah Webster (1758-1843) published the first volume of A Grammatical Institute of the English Language – known in later editions as the American Spelling Book, then the Elementary Spelling Book, or ‘the Blue-back Speller’ – which conveyed to generations of Americans Webster’s nationalistic and pedagogical enthusiasm for spelling reform and simplification. Still in print, it was clearly the most successful language arts text ever published. In 1806 he published his Compendious Dictionary of the English Language, which included a generous sampling of his reformed spellings. In 1828 far fewer of his reforms were included in his much expanded An American Dictionary of the English Language.

Webster competed with and could be acidly critical of Samuel Johnson, whose A Dictionary of the English Language (1755) dominated spelling practices not only in Britain but also in the 18th and early 19th centuries in America. Webster argued for an American spelling distinct from British for reasons partly pedagogic, partly nationalistic, partly economic. In time American readers, writers, and publishers began to be influenced by the finally rather modest spelling reforms in Webster’s later spellers and his 1828 dictionary, and by the fact that Americans tended to prefer simplification and were less concerned than Johnson had been with British traditions. For more on Webster (and Johnson), see my “The Evolution of British and American Spelling” in The Routledge Handbook of the English Writing System (V. Cook and D. Ryan, eds.) London & New York: Routledge, 2016. pp. 275-292).

Wessex and the West Saxon dialect. See Alfred the Great.