Word choice

Last modified: 2026-11-13 12:24:54 (PST)

> **AI-Assisted Authorship**
>
> This chapter was written with the assistance of GitHub Copilot, which expanded on outlined ideas and draft notes provided by the author. The content represents the author's perspective and has been reviewed for accuracy, but the detailed prose was generated through AI assistance.

I recommend trying to replace Latin-derived words and phrases with Old English-derived equivalents ("Anglish" words) where possible; it generally makes writing simpler and easier to read. Latin words create artificial barriers to understanding. Many Latin-derived words commonly used in scientific writing are composed from roots and affixes which are not commonly used in their basic forms; hence, readers cannot determine the meanings of these words by decomposing them. Instead, they need to memorize the meanings of these words directly. In contrast, the components of composite Anglish words are typically also used individually, so the meanings of the composites can be derived directly. Table 1 lists some common Latinate words and phrases and Anglish alternatives.

Table 1: Commonly used Latin words and phrases and Anglish alternatives

Latin	Anglish
prior to	before
necessary	needed

See also https://bark-fa.github.io/Anglish-Translator/

English has three layers of vocabulary because of its history (Wikipedia contributors 2026):

- Old English came to Britain with Germanic-speaking settlers in the mid-5th to 7th centuries. Its words are the source of the Anglish alternatives above.

- French and Norman words entered after the Norman Conquest of 1066, about 10,000 of them in Middle English, particularly terms for government, church, law, the military, fashion, and food.
- Latin and Greek words entered in the Renaissance, as increased literacy and travel helped writers adopt many foreign words.

Because the borrowed words came in through government, church, law, and learned writing, many of them still read as formal and learned. Their Old English counterparts read as plain and everyday. A reader does not notice this contrast in a single word, but across a paragraph it sets the tone: "commence" and "utilize" sound official, while "begin" and "use" sound direct. When you choose between two synonyms, choose the one whose tone matches your readers and your purpose.

To find a replacement for a word, look the word up on the Open English Wordnet website (source on GitHub). Open English Wordnet groups English words into sets of synonyms, called synsets. A word with several meanings appears in several synsets, one for each meaning, so choose the synset that matches the meaning you intend. Then pick the word in that synset that your readers are most likely to understand easily.

Overused vocabulary lists more words, many of them Latin-derived, that writers reach for by reflex, with plainer alternatives.

I am aware that this book, and even this chapter, contains many Latin word choices where there are Anglish alternatives. It is a work in progress, and also, I am not advocating 100% Anglish purity. Use whichever words and phrases you think your readers are most likely to understand easily. Preferring Anglish is merely a useful heuristic to help achieve our ultimate goal of producing clear, easy-to-read writing.

Just to be clear, although I prefer Anglish words, I have no particular preference for Anglish people or culture; it is only a practical consideration, based on the realities of English as the current default language of science and the relatively-recent hybridization of the English language.

Avoid vague and metaphorical language

Write simple, direct prose (Ezra, 2026-11-03). Scientific claims should be precise enough to test. Vague or metaphorical phrases can sound evocative, but they leave the reader guessing at the specific quantity or mechanism you mean. Whenever you can, replace a metaphor with the precise term for the thing you are describing. Test every word, whatever its part of speech: it must be literally true of what it names or describes. A point does not "sit" at a height, and an estimate does not "live" in an interval. Write that the point is at the height and the estimate is in the interval. No list of phrases is complete, so apply the test to every word instead of looking a phrase up. Table 2 lists some common vague phrases and more precise alternatives.

Table 2: Vague or metaphorical phrases and more precise alternatives

Vague or metaphorical	More precise
infection pressure	incidence rate
disease burden	incidence rate, or prevalence proportion
the signal was strong	the effect size was large
a huge effect	a risk ratio of 4.2

The precise alternative is not always shorter, but it tells the reader exactly which quantity you measured, which makes your claim verifiable. When a metaphor genuinely aids intuition, state the precise quantity first, then offer the metaphor as a secondary aid.

Informal idioms are metaphors too. "Kicks in", "the whole point" and "getting there" stand in for a literal claim; state the claim. Use literal language explains why idioms exclude some readers, and gives more examples. Stock metaphors such as "pave the way" and "shed light on" are clichés, and Overused vocabulary lists more of them.

Money words are idioms too when no money is involved. "Costs us", "buys us", "at the price of", "pays off", "affordable", "cheap" and "the whole bargain" borrow from commerce to describe a trade-off, and the reader must translate them back. A verb "trade" does the same: "that trade" and "we trade accuracy for speed" put a bargain where there is a relationship between two quantities. Say what is lost or gained, and which quantity changes with which. Keep a literal use: a computational cost, a monetary cost, or a cost function that the text defines as a term.

✗ Nesting the cross-validation loop costs another factor of K in computing time.

✓ Nesting the cross-validation loop multiplies the computing time by K.

✗ Continuity buys us the intermediate value theorem.

✓ Continuity gives us the intermediate value theorem.

✗ The penalty costs us a little flexibility at the ends of the curve.

✓ The penalty reduces the curve's flexibility a little at its ends.

✗ That trade is the whole bargain of gradient descent.

✓ Gradient descent decreases the objective function only when each step is small enough for the linear approximation to hold.

Do not give a mathematical object a mind. A derivative does not "recommend" a direction, a model does not "know" or "want" anything, and an approximation is not "trustworthy". Each such phrase hides a precise claim, and the reader needs that claim stated plainly. State

the condition under which the claim holds, with the inequality or the order of the error, and display it when it is a key equation.

> ✗ The derivative recommends a step, and the recommendation is trustworthy only as far as the step is small.
>
> ✓ The linear approximation has error $o(\|\underset{\sim}{h}\|)$ as $\underset{\sim}{h} \to \vec{0}$, so the objective function decreases along the negative gradient only for steps $\underset{\sim}{h}$ that are small enough.

Be as specific as feasible. When the text has a specific object, name it, and do not let a generic placeholder stand in for it, such as "some set", "a set A" or "the object". A placeholder makes the reader ask which set is meant, and the writer has already answered that question. Keep a placeholder only where the statement is abstract on purpose, such as a general definition that holds for any set. The same aim, stated from the other side: state a result at the generality it holds, no narrower than its argument needs. A fact that a proof proves only for one case, when the argument works for all cases, is a lemma left unextracted (Extract a lemma wherever feasible).

> ✗ The mode is an element of the set A at which the density is largest.
>
> ✓ The mode is an element of the range of the random variable at which the density is largest.

The `check-avoided-words` check in gha warns on a domain, codomain, range or sample space named by a bare capital letter or by "some", such as "domain A" or "some range". It cannot tell whether a specific object exists, so a warning only asks you to check.

Some single words cause a different problem. Words such as "obviously", "clearly" and "trivially" can assert that a claim needs no support instead of supplying any. Do not substitute flippancy for support says when those words are a problem and what to write instead.

Follow an abstract statement with an example

An abstract or general statement is fine, as long as a concrete example follows it. A vague statement is not: one that names no specific case leaves the reader to invent one, and they may invent the wrong one.

For example, consider this bullet from a list of what machine learning needs:

> statistics to say what a finite sample of data supports

The bullet is general, which is fine. What it lacks is an instance that shows the reader what "supports" means. Adding one turns it into something a reader can check:

statistics to say what a finite sample of data supports: if most of the spam in a pile of training emails mentions money, statistics tells us whether that pattern should hold for next month's emails or is a quirk of this particular pile

When you write a general claim, ask what one specific case of it looks like, and put that case right after the claim.

Say which thing

The word *thing* almost never survives a second look. Name the object, quantity, step or idea it stands for.

✗ The main thing to check is the residuals.

✓ Check the residuals first.

The plural and the filler phrases built on it ("a number of things", "other things", "things like that") are the same problem. Each one tells the reader that something exists without saying what.

✗ Several things can inflate the standard errors.

✓ Correlated predictors and clustered observations can inflate the standard errors.

If you cannot name the object, quantity, step or idea, the sentence is not yet ready to write.

Give a technical word one meaning

A word that names a technical concept keeps that one meaning on every page. Two fields often use the same word for different concepts. When a reader encounters the word in its second sense, they must stop and work out which concept the sentence means.

Pick the word's established meaning in the course sequence, and name the other concept with a different phrase. If the other field's word is still useful, list it as a labeled alias in the definition of the concept it names.

For example, *hypothesis* means a statistical hypothesis: a claim about a population, such as a null hypothesis that a test examines. In machine learning, *hypothesis* also names a candidate prediction function.

✗ The learner picks one hypothesis from the hypothesis set.

✓ The learner picks one function from the prediction function family.

The definition of the prediction function family can then say that machine-learning texts call it a *hypothesis class* or *hypothesis space*.

Call the distribution Gaussian

The word *normal* names several unrelated concepts in mathematics and statistics. It names the normal equations of least squares, a normal vector perpendicular to a surface, and a normal subgroup. In everyday use, it also means *typical*. When the subject is the distribution, *Gaussian* has no such competing meaning in the same context, so use it.

> ✗ The errors need not be normal.
>
> ✓ The errors need not be Gaussian.

Use the same word in every phrase built on the distribution:

- Write *Gaussian distribution*, not *normal distribution*.
- Write *standard Gaussian*, not *standard normal*.
- Write *multivariate Gaussian*, not *multivariate normal*.
- Write *the errors are Gaussian*, not *the errors are normally distributed*.

Keep *normal* where it names a different concept, as in the *normal equations*. Keep *normal* in a quoted title or a cited name, such as `rnorm` in R, where you report what the source calls it.

Call the exponentiated coefficient an odds multiplier

In a logistic regression model, a change $\Delta\underset{\sim}{x}$ in the predictors multiplies the odds by $e^{\vec{\beta}\cdot\Delta\underset{\sim}{x}}$, and a one-unit increase in predictor j multiplies them by e^{β_j}. Call that number the *odds multiplier* of the change, not the *odds factor* (Ezra, 2026-11-12). *Multiplier* says what the number does to the odds, while *factor* already has other meanings in these notes:

- a categorical variable in R;
- a latent variable in factor analysis;
- a number that divides another.

> ✗ The odds factor of a one-unit increase in age is $e^{\beta_{\text{age}}}$.
>
> ✓ The odds multiplier of a one-unit increase in age is $e^{\beta_{\text{age}}}$.

The symbol stays θ_{ω}, written with the shared `\oddsfactor` macro: the rule is about the word in the prose, not the macro's name.

Name a quantity for what it is

A name should say what kind of thing the quantity is. Each kind of quantity has its own name:

- **A variance** measures spread.
- **A risk** is an expected loss.
- **An error** is a realized difference between a prediction and an outcome.

Calling a variance an error, or calling a quantity "irreducible" without saying given what, misleads the reader.

For example, the variance of the outcome given the inputs is the *residual variance*. It is a component of the expected squared prediction error, not an error, and it is irreducible only given the inputs the model uses: a further input that predicts the outcome reduces it.

> ✗ The irreducible error is the part no model can remove.
>
> ✓ The residual variance is the part no model using the same inputs can remove.

Mention the common name once, as an alias in the definition, and use the accurate term everywhere else. No mechanical check can tell a misleading name from an accurate one, so this rule is enforced by review.

Write the full phrase

Do not leave out words the phrase needs to be complete. A clipped phrase such as "the true mean function" asks the reader to supply the missing words, and the reader may supply different ones than you meant. Readers who are learning English, and readers who read literally, cannot always recover them. Name what the object is a function of, and what it is conditional on.

Name a function with the word "function": write "cost function", "loss function", "logit function", and "logistic function", not "the cost", "the loss", "the logit", or "the logistic", when the sentence is about the function itself. Do not clip a technical term, even one that the text has introduced. Write the full term at every use. An acronym or initialism is neither clipped nor elliptical, and this rule allows it. The rule covers every named function. A bare name without "function" means the value of the function at a specific input, and "X function" means the function itself:

- **Cost:** the cost of a particular parameter vector is the value of the cost function there.
- **Loss:** the loss of one observation is the value of the loss function at that observation.
- **Logit:** the logit of a particular probability is the value of the logit function at that probability.

Name the input in the sentence or its context, so the reader can tell a value from the function.

The same rule applies to symbols. Write a function with its argument, such as $\mu(x)$, not the bare function name, unless the sentence is about the function as a whole and the page has said so. A symbol in a sentence should look like the same symbol in the display equation it refers to.

Example 0.1 (Writing the full phrase and the argument)**.** The display equation is $Y = \mu(x) + \varepsilon$.

✗ The outcome variable is the true mean function plus noise.

✓ The outcome variable is the true mean function of the outcome variable given the input value x, namely $\mu(x)$, plus noise.

Example 0.2 (Naming a function in full)**.**

✗ The cost is convex in the parameters.

✓ The cost function is convex in the parameters.

✗ The logit is the inverse of the logistic.

✓ The logit function is the inverse of the logistic function.

✓ The cost of the initial parameter vector is 4.2, and each step lowers it. (A value of the cost function at a specific input, so "function" is not needed.)

Choose each preposition on purpose

A preposition states a relationship between two things. Small words such as *of*, *for*, *with* and *using* are easy to reach for by reflex, and each one can stand for many different relationships (ownership, purpose, cause, part and whole, tool, source). When the sentence does not say which relationship it means, the reader has to guess. Readers who are learning English, and readers who read literally, guess worst.

For each preposition, ask what relationship it states. If you cannot answer in one phrase, replace it with a verb that names the relationship.

Example 0.1 (Naming the relationship)**.**

✗ The error of the prediction is large.

✓ The prediction misses the observed outcome by a large amount.

Of could mean the error the prediction makes, the error in the way it was computed, or an error that belongs to the prediction in some other way. The verb *misses* says which.

> ✗ We fit the model for the training set.
>
> ✓ We estimate the model's coefficients from the training set.

For could mean on behalf of, in order to serve, or using. *From* states the relationship exactly: the training set is the source of the estimates.

Keep a plain preposition when it states the relationship exactly, such as *in* for location and *by* for the agent of a passive verb.

Refer to a place in the document by its label

Do not use "above" or "below" for a position in a document. The same source renders as a web page, a slide deck and a handout, and each format orders and breaks the content differently, so "above" can point at nothing, or at the wrong thing. Style guides advise "earlier" or "preceding" and "later" or "following" for a position in a document (Google, n.d., entry "above"; Google, n.d., entry "below"). Prefer the label of the result, figure, table or section, because a label tells the reader what to find and survives a reordering.

Example 0.1 (Replacing "above" with a label)**.**

> ✗ The next figure draws the fits behind the table above.
>
> ✓ The next figure draws the fits behind `@tbl-fits`.

The rewrite needs a label on the table, so labeling it is part of the fix. Quarto renders `@tbl-fits` as "Table 2", and the number follows the table when the page changes. A comparison, such as "a value above the threshold", does not refer to a position in the document and is fine.

Leave out words that judge difficulty or add nothing

"Just" and "easy" ("easily") rarely carry meaning. The Google style guide says "just" is "a filler word that you can delete without affecting your meaning" and says to try eliminating "easy" and "easily" "because usually the same meaning can be conveyed without" them (Google, n.d., entries "just" and "easy, easily"). "Easy" also judges the reader: a reader who finds the step hard is told the difficulty is their own. Do not substitute flippancy for support covers the same problem for "obvious" and "clearly".

Keep the word when it carries meaning. “Just before the loop ends” states a position in time, and deleting “just” changes it.

Give a date or a version instead of a time-relative word

Do not write “now”, “new” or “currently” for a state that will change. The Google style guide says these words “can render such documentation inaccurate, outdated, or unmeaningful” (Google, n.d., “Timeless documentation”) and says of “currently”: “Avoid because this word is implied” (Google, n.d., entry “currently”). Say which date, version or release the statement holds for.

The guide does not list “recently”, but it fails the same way. In course notes it also fails when pages reorder, because “recently” assumed an order of lectures that the page does not fix.

Example 0.1 (Replacing “recently” with a pointer)**.**

> ✗ A missing *that* is a common cause of garden-path sentences, as we recently saw.
>
> ✓ A missing *that* is a common cause of garden-path sentences, as Avoid garden-path sentences shows.

Keep “this” and “also” out of consecutive sentences

The Stanford course notes on mathematical writing say to “avoid words like ‘this’ or ‘also’ in consecutive sentences”. They explain that such words, and “unusual or polysyllabic utterances”, “tend to stick in a reader’s mind longer than other words, and good style will keep ‘sticky’ words spaced well apart” (Knuth et al. 1987, point 9). So avoid using “this” or “also” in two sentences in a row.

Example 0.1 (Spacing out “this” and “also”)**.**

> ✗ This model also fits the training data well. This fit also holds on the test set.
>
> ✓ The model also fits the training data well. Its fit carries over to the test set.

The ✗ pair uses “this” and “also” in both sentences. The ✓ pair rewords the sentences but keeps both claims, and uses “also” once and “this” not at all, without a synonym such as “too” in the second sentence.

The same point of the notes says to "use parallelism when parallel concepts are being discussed" (Knuth et al. 1987, point 9). A word repeated to make two statements parallel is repeated on purpose, so keep it, as Repeat the wording of parallel results, then point out the difference describes.

Minimize unnecessary jargon

Jargon is vocabulary specific to a field. Some jargon is necessary: a precise technical term can replace a long explanation, and readers in the field expect it. But jargon becomes a barrier when a normal word would work just as well, or when you use a common word with a special, field-specific meaning without defining it.

Prefer normal words used in their usual sense. When you do need a technical term, define it clearly at first use, then keep using the same term rather than switching between synonyms. When the field itself uses several names for the concept, list them where you define it (see Name the synonyms).

Example 0.1 (Replacing unnecessary jargon)**.**

✗ We leveraged a bespoke pipeline to operationalize the ingestion of heterogeneous data modalities.

✓ We wrote a custom program to combine data from several sources.

The plain version is shorter, and it states what was actually done.

For each technical word, ask whether it carries meaning that a plainer word would lose. If it does not, replace it with the plainer word. If it does carry necessary meaning but a reader outside your subfield would not understand it, define it clearly at first use.

Google. n.d. *Google Developer Documentation Style Guide*. Web page. https://developers.google.com/style.

Knuth, Donald E., Tracy Larrabee, and Paul M. Roberts. 1987. *Mathematical Writing*. https://jmlr.csail.mit.edu/reviewing-papers/knuth_mathematical_writing.pdf.

Wikipedia contributors. 2026. *History of English — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/History_of_English.