Duke Graduate School Scientific Writing Resource

Lesson 2: Cohesion, Coherence, and Emphasis

In lesson 1, we covered how the structure of the sentence cues the reader to important information. If you structure your sentences carefully, you encourage readers to interpret your meaning correctly. Structure has several other important uses: First, you can structure sentences for emphasis, drawing attention to the most important part of the sentence. Second, you can structure your sentences and paragraphs to flow — that's what cohesion and coherence are all about.

Cohesion is the degree to which sentences "glue." Coherence is the logical division of the writing into internally consistent units (usually paragraph units). In this lesson, you will learn ways to improve the cohesion and coherence of your writing.

Principles

1. Put new information last
2. Use passive voice judiciously
3. Make sure the first and last sentences of a paragraph match

Principle 1: Put new information last

Ideas or characters that have not yet appeared in your manuscript are called New information. New means unfamiliar. "Old information" is something familiar to the reader, either because it's background knowledge or because you've already introduced it. Your sentences will contain both new and old information — think carefully about where you put them. Most readers will find your writing more clear if you consistently begin sentences with familiar (old) information and conclude sentences with unfamiliar (new) information.

What happens when you begin a sentence with new information? Your reader gets a new idea without any context. He or she may try (incorrectly) to link this information to the previous sentence. After reading the rest of the sentence, the reader may have to revise his or her understanding. If you do this too much, it makes your writing confusing because it lacks cohesion. Going backwards like this slows the reader down and takes energy. Beginning sentences with old information makes writing cohesive. It also allows you to put new, important information in the position of emphasis at the end of the sentence.

Imagine these sentences in an article about farming:

> Farmers try to provide optimal growing conditions for crops by using soil additives to adjust soil pH. Garden lime, or agricultural limestone, is made from pulverized chalk, and can be used to raise the pH of the soil. Clay soil, which is naturally acidic, often requires addition of agricultural lime.

About This Guide

This guide to scientific writing was originally created in 2010-2011 by Nathan Sheffield for the Duke University Graduate School and funded by a Duke University Graduate School Teaching mini-grant. This current site is maintained by the Duke Graduate School. If you have questions about this site, please contact gradschool@duke.edu.

It is difficult to see at first, but the second and third sentences have the same problem: they begin with new information. If we separate the sentences and color the old information and the new information it becomes easier to notice:

> Farmers try to provide optimal growing conditions for crops by using soil additives to adjust soil pH.
>
> Garden lime, or agricultural limestone, is made from pulverized chalk, and can be used to raise the pH of the soil.
>
> Clay soil, which is naturally acidic, often requires addition of agricultural lime.

Now, let's follow the reader through this paragraph. When the reader begins sentence #2, reading "Garden lime…", there is little context; the reader may guess "limes" are a crop we will now discuss, or that it is a "soil additive". There are at least two possible connections to the previous sentence, and readers will be split. At the end of the sentence, we are given the context and the connection: "raise the pH". This backward-glance at the end of the sentence causes the reader to backtrack, costing concentration.

The third sentence is also problematic. It begins with "Clay soil…", similarly without context. The reader may then think "clay soil" as another additive, perhaps one that lowers the pH? At the end of the sentence (requires…lime), you finally get the connection back to the previous sentence and the context for "clay soil," but this causes the reader to backtrack.

To solve the problem, we can try swapping the new and old info. Here's one possible revision:

> Farmers try to provide optimal growing conditions for crops by using soil additives to adjust soil pH. One way to raise the pH of the soil is an additive made from pulverized chalk called garden lime or agricultural limestone. Agricultural limestone is often added to naturally acidic soils, such as clay soil.

In the revision, each sentence leans forward to new information at the end, instead of tying backward at the end. This makes the sentences easier to read, because the reader doesn't need to jump around in thought process.

When your sentences "glue", your writing is said to be cohesive. If your sentences are regularly beginning with unfamiliar concepts, your writing won't be very cohesive. This is where the passive can be so useful: if it improves cohesion. More on that in the next principle.

Putting new information last also helps with emphasis: readers naturally emphasize the ideas at the end of the sentence. Putting the new, important information at the end will help inform the readers of what you intend to emphasize.

Revision Technique

Read through your manuscript carefully. In each sentence, underline any pieces of new information (unfamiliar to the reader at this point in the manuscript). Make sure your sentences begin with an appropriate backwards link, and not with an unfamiliar concept.

Principle 2: Use passive voice judiciously

Sentences are in passive voice when the subject in the sentence is the object of the action. For example:

Active: The dog chased the ball.

Passive: The ball was chased by the dog.

Passive: The ball was chased.

Using passive can have two important consequences:

1. The order of the subject and verb are switched. (First passive example above)

Active: Dog → Ball
Passive: Ball → Dog

2. The doer of the action can be omitted. (Second passive example above)

Active: Dog → Ball
Passive: Ball → ???

Passive voice isn't inherently bad. It can actually be quite useful. The problem is that some writers incorrectly think passive voice is inherently scientific. In fact, some students are taught that passive voice is more objective. Really, the way you write doesn't make your experiments any more objective; instead, your results should speak for themselves.

For whatever reason, many scientists rely on passive voice excessively. But scientific journals would rather you use active voice:

> Nature journals like authors to write in the active voice…-Nature

> Choose the active voice more often than you choose the passive… -Science

Why? Here are some possible consequences of relying on passive voice:

1. Ambiguous characters

A consequence of passive voice is that the actor can be omitted, which is common in scientific writing. Sometimes this makes sense, other times is causes confusion. It's OK to omit the actor if it is self explanatory, understood, or unimportant. For example, passive voice can be effectively used in a methods section to focus the reader on the method (instead of on the actor). It is not OK to omit the actor if there are multiple possibilities, leaving your reader to guess. For example:

> The DNA was sequenced using the n-terminus method (Smith et al. 2004).

In this example, who sequenced the DNA? Is the paper being cited because Smith et al. did the sequencing, or because they invented the n-terminus method? Any time you leave multiple possibilities, you divide your readers. Some readers will misinterpret your intent.

2. Dangling modifiers

When you write passive sentences, be careful not to dangle your modifiers! From an editorial in Science:

> Our chief objection to the passive voice is that it sometimes seems to make authors forget to watch for dangling modifiers.

A dangling modifier is a modifying phrase whose implicit subject does not match the explicit subject of the clause it modifies.
Dangling modifiers are common errors in scientific writing. If you want to learn more, I wrote a focus article on dangling modifiers.

3. Wordiness

All else being equal, shorter writing is better: it takes less time to read and it uses less space. These are important things to consider in scientific writing. Readers benefit from less reading (it takes less time), and scientists are also regularly subject to journal space constraints. Whatever else is true of passive voice, it is a fact that passive voice tends to increase length (however slightly). When every word counts, active voice can help keep writing concise.

Advantage of passive voice

The key use of passive voice is that it switches the order of a sentence. This is hugely important in light of principle 1 in this lesson. Use passive voice when it moves the old information to the front and new information to the back (see Principle 1). Guide your writing with the rule “Put new information last” instead of the rule “Always use passive voice.” Use the passive as needed to keep the flow, and always provide the actors if there is a possibility of confusion.

The point of this principle is not to eliminate passive voice, but to increase your awareness. Choose passive voice for a reason, not because you think it “sounds scientific.” There’s a lot more to be said about passive voice. If you’re interested in a more in-depth treatment of the active/passive voice discussion, I wrote a focus article on passive voice in science writing.

Revision Technique

When you revise for new/old information placement, use active/passive switching to keep the placement of information consistent with expected structure.

Principle 3: Make sure the first and last sentences of a paragraph match

This principle is called coherence. Usually, when readers refer to the “flow” of writing, they are referring either to coherence, or to cohesion (Principle 1). When writing is coherent, it stays on topic in expected units. Readers usually expect thoughts to be expressed in paragraph units. A single paragraph corresponds to a single thought. Each sentence in the paragraph should support that main point.

Just because your sentences stick together by including appropriate backwards links, it doesn’t mean your writing is coherent. Here is an example of a paragraph that is cohesive, but lacks coherence:

> My favorite animal is the domestic cat. Cats were domesticated almost 10,000 years ago in ancient Mesopotamia. Mesopotamia is a name that literally means “the land between two rivers,” taken from Greek. The Greek language is one of the oldest written languages, and its alphabet forms the basis of many other writing systems, including Latin. Latin …

This paragraph has great cohesion (one sentence leads to another), but it wanders through topics. Coherence is a paragraph-level principle (which makes it difficult to provide examples for).

Revision Technique

Test for coherence:
Read the first and last parts of each paragraph. Do the topics match? To be more thorough, make sure each sentence in a paragraph supports the main point of that paragraph.

Examples

Example 1

At the beginning of a paper wanting to emphasize detecting positive selection:

> Detecting positive Darwinian selection at the DNA sequence level has been a subject of considerable interest.

Readers naturally emphasize the end of sentences. What do we want to emphasize? Probably it would be better to emphasize either “positive selection” or “DNA sequence level”, depending on the point of the paper. Here are some possibilities:

> One subject of considerable interest has been detecting positive Darwinian selection at the DNA sequence level.

> One subject of considerable interest at the DNA sequence level has been detecting positive Darwinian selection.

The first revision seems better in a discussion of positive selection that wants to emphasize testing at the DNA sequence level (as opposed to testing at the protein level, or some other level). The second revision seems suited to a discussion of several characteristics of DNA, of which one to emphasize is positive Darwinian selection.

Example 2

> Improvements are expected in the predictive power of all the scores being computed on multispecies alignments.

This example has two passive-like verbs: are expected and being computed. If we try to eliminate those, we might get a more direct revision:

> Our method will improve the predictive power of all multispecies alignment scores.

Example 3

> A survey is given of differential expression analyses using the linear

> modeling features of the package.

I believe this sentence is the result of a reliance on passive voice. The phrase survey is given of is difficult to parse. It gets more difficult because the using phrase can refer either to survey or analyses. It's not immediately apparent which the author means. Depending on what the author meant to say, we could revise like so:

> We use the linear modeling features of the package to survey differential expression analyses.

> We survey differential expression analyses that use the linear modeling features of the package.

Example 4

> Using sarkosyl to induce nuclear run-on, the transcriptionally inactive b-globin gene in mature erythrocytes was demonstrated to harbor high levels of Pol II at 5′ proximal regulatory regions.

This example relies on passive voice and ends up with a dangling modifier. If we simply eliminate the passive in the main clause, we can get a revision like this:

> Using sarkosyl to induce nuclear run-on, Smith et al. showed that the transcriptionally inactive b-globin gene in mature erythrocytes harbors high levels of Pol II at 5′ proximal regulatory regions.

Example 5

> We identified genes that are differentially expressed between species. A phylogenetic tree based on the number of differentially expressed genes between species recapitulates their known phylogeny.

The second sentence begins with "a phylogenetic tree", which is new information in this context, while "differentially expressed genes" is old information. Also, because of the complex subject, most of the words in the second sentence are between the subject and verb:

> A phylogenetic tree based on the number of differentially expressed genes between species recapitulates their known phylogeny.

Let's try to simplify the subject, bring the simple subject closer to the verb, and put new information last:

> We identified genes that are differentially expressed between species. The number of differentially expressed genes can be used to build a phylogenetic tree that recapitulates the known phylogeny.

In the revision, the second sentence has now a backwards link to the previous sentence (old info.), instead of starting with a new concept. In addition, the subject is less complex, and the subject-verb distance has been reduced (phylogenetic tree is now near recapitulates).

Worksheet

lesson2.pdf

Continue to lesson 3