How to Do Great Research

Creativity and grad school survival advice from Professor Nick Feamster

Storytelling 101: Writing Tips for Academics

Posted: October 11, 2013 | **Author:** Nick Feamster | **Filed under:** advice, writing | **1 Comment »**

Little did I know that a Ph.D. in computer science would teach me more about how to write than any other previous experience I'd had. Indeed, research isn't only about making new discoveries, but also about *expressing* those discoveries so that other people can appreciate them and their significance.

Below, I'll present some thoughts on writing that I've picked up over the years and continually aim (and encourage students) to put into practice. I'll focus on story-telling, presentation, and efficiency: how to write your paper so that it clearly, cleanly, and efficiently tells a good story.

What this post is not about. At least in this post, I won't write about grammar—there is too much English grammar to cover in a single post, and there are plenty of good courses one can take to learn grammar. I also won't talk about style (common phrases) or usage (e.g., commonly misused words). For now, my only advice on the topic of learning English grammar and style—both for native speakers and foreigners—is to read daily. Reading the *New York Times,* or *The Economist* are great places to start: the writing is generally of high quality, and you'll also have something interesting to talk about outside of the lab. I have the impression that foreign students can tend towards reading their local news from home; that's understandable, but I would encourage those students to *also* read periodicals and publications that are written in the style of English that people are used to reading in the Western world (indeed, English in different parts of the world has different tones, styles, idioms, and turns of phrase).

Why Does Writing Well Matter?

Writing is part and parcel of the research process. Writing communicates your research results, and—perhaps contrary to intuition—writing well is even more important when the results that one wants to present are "strong" (after all, if the results are that great, you should want as many people as possible to benefit from understanding them). But writing isn't just something that happens at when the research is "done"; rather, *writing should occur throughout the course of a research project*. Writing and thinking are tightly coupled—writing is a reflection of our own thoughts, and writing well can actually sometimes clarify our own thinking about a problem and lead us to a solution.

Communicating your results. A research finding or discovery that others cannot understand or appreciate is as good as no discovery in the first place. A well-written paper with a half-baked idea that people can understand and appreciate will often be accepted to a conference over a paper with a (possibly) stronger result that nobody can understand or appreciate. Cynics may dismiss this as "unfair", as "laziness" on the reviewer's part for not reading the details thoroughly, or as "accepting a paper because it was well-written" (insinuating that the results themselves might not be that strong). The fact of the matter is that none of this matters: Independently of the quality of the research, the results must be written well, or people will not be able to appreciate the value of your work. Writing well is perhaps one of the most important skills that budding researchers must attain; it is equally important as learning how to conduct the research.

Clarifying your own thinking. Writing is a reflection of your thinking. You should never be content with vague language or imprecise descriptions. *Muddy writing often reflects muddy thinking.* An attempt to write down a concept, idea, or solution can actually help you realize that you don't know what you are talking about! (And, help you crystallize your thinking.) You may have had the experience of realizing you don't understand something as well as you thought you did when someone asks you to explain something and you have trouble. Writing achieves the same effect: as you attempt to write something down, you may realize that you don't know why you are working on something, why a particular phenomenon appears, and so forth. If you don't force yourself to *try* to write down your explanations and observations, you may never discover that you don't know what you're doing until it's too late. Therefore, *start writing early* in the course of a research project. Most of the early text you write may not find its way into the final version of the paper, but taking the time to express your ideas on paper can help you crystallize your thoughts and guide your research.

Tell Me a Story

Every paper should tell a story. For a given research result, there are many stories that a paper *could* tell, and it's important to think about what the story should be. This decision should be made early on in the writing process. The story can be changed later on, to match the strongest results in the paper, or if the context for the research changes (for example, sometimes when working on a research problem, people can find it necessary to re-formulate the problem that they are solving). But, it's always good to figure out a story before you go too far down the path.

Determining the Big Picture: Composition and "Flow". *Composition* is the structure and flow of your paper writeup, and how individual paragraphs fit together. When people talk about the "flow" of a piece of writing, they're talking about composition: how sentences flow together to form paragraphs, how paragraphs flow together to form sections, and how sections flow together to form a paper. *Composition (or "flow") is the most important aspect of technical writing.* I've found that reviewers (and readers) are often forgiving of the occasional grammatical error, especially if a paper was written by non-native English speakers. The human brain is pretty good at applying "error correction" on grammar. Correcting "flow", on the other hand, is extremely hard for a reader to do. Often, the reader will complain that they "can't understand where [the paper] is going" and simply give up. Grammar is also easier to teach (and learn) than composition because there are fixed rules for grammar; composition is more of an art form, and it does involve a bit of intuition. It involves thinking about the big picture, the story that you want to tell, and how the pieces fit together. It takes practice to develop intuition for composition, and I've found that native English speakers are not necessarily any better at composition than foreigners are (that's good news for Ph.D. students coming from abroad: the playing field is actually more level than it appears). Even though composition is a bit of an art form, I've found that the following tips can be incredibly helpful for improving flow:

- **Build the scaffolding before filling in the details. Write sections and topic sentences first.** There is always the temptation to simply jump in and start writing full paragraphs. This rarely ends well. Nobody would build a house before first drawing plans. After the plans comes a foundation, and scaffolding. Only after the structure is in place can many of the details be filled in. You should think about writing a paper in the same way: before you can fill in the details, you need to figure out how you want the paper to read when its done; this is your outline. Start at a high level, by outlining sections. Once you have an outline of sections, you can begin to slowly fill in more details: outline each paragraph within a section by writing only the topic sentence for the paragraph. See whether your story makes sense if you read each topic sentence in sequence. Your "paper" should roughly make sense if you read these topic sentences in sequence (in fact, some reviewers or readers may try to read your paper in exactly this fashion, by skimming topic sentences!). If reading your topic sentences in sequence doesn't make sense, something is wrong with your flow. Determining which topic sentences to write takes a bit of practice, but I find that making a bulleted list of points that you want to include in the section and then clustering those points into related points and topics can help identify the topics you want to cover (and, hence, the topic sentences you want to write). If you find that the topics you want to cover are not all related to one another, you may need to create additional subsections. Once you have topic sentences, you can fill in paragraphs, which are groups of closely (and logically) related sentences. If your sentences aren't closely related, they don't belong in the same paragraph and you may need to consider adding a new topic (and topic sentence).
- **Assess balance and budget pages accordingly.** Once you have your outline, it's important to budget pages, to figure out how much you plan to write about each topic. Topics of equal importance or equal relevance should receive "balanced" coverage in the paper. If your system has two components of relatively equal importance, then the sections covering those components should be of approximately equal length in the paper. Once you have your outline (from the previous step), you can assess how much space you want to budget for each point. If some points in a section are more important than others (and hence require more page budget), you can go back and think about how to re-balance your

writeup (e.g., by creating a new devoted section to the topic that's more important, shortening or eliminating the less important point, and so forth).

- **Follow convention.** Continuing the analogy of a house or building, many designs follow convention: a restroom in a restaurant is often in the back. A powder room in a house is often near the front door. These common "design patterns" make life easier for us when we visit a new building or house. Similarly, your reader is a visitor of your thought process. A reader who has read (and written) many research papers will similarly expect the structure of your paper to follow certain conventions. These conventions differ by field: some fields write papers completely in the passive voice, whereas others use active voice. Some fields prefer related work sections at the beginning of the paper, and others prefer related work at the end. Some fields expect an introduction that spends time framing the problem, and others expect an introduction that starts with a clear statement of the main result. There may be an expectation of (or aversion) to mathematics in the main line of the paper. Within a section, people may expect the introduction to flow in a particular way (in computer science, Jim Kurose's **description of how to write an intro** is the best example I've seen for writing a conventional paper introduction).

- **Signpost your paper.** Even if your paper plan and outline make perfect sense to you, the reader doesn't start out "on the same page" as you. Even if the builders have built a house completely according to the blueprints, even if the design of the house follows convention, visitors still typically appreciate an explanation of where the bathroom is. The reader needs instructions for how to read the paper and a clear view for how the paper and story will proceed. Signposts throughout the paper help the reader understand how the paper is organized. Examples of signposts include: an outline of the paper at the end of the introduction ("The rest of this paper proceeds as follows."), a preamble to each section ("In this section, we discuss…"), declarative subsection titles, and (within reason) bold paragraph headings (such as those in this blog post).

Make a Good Impression

A reader forms an impression about your research based on your presentation of the work. Just as people quickly form first impressions about other people that may be difficult to subsequently correct, your reader will quickly form a first impression of your work based on a quick read of the title, abstract, and introduction and perhaps a quick skim through other parts of the paper. Therefore, it is important to make a good impression by writing a clear abstract and introduction (and also ensuring that the rest of the paper also is commensurate with the impression that you create).

Spend a lot of time on your introduction; start early. The introduction summarizes the story of your paper (now, it should be doubly clear why you need a story!) and is the most important part of your paper. Many people won't read past the introduction, even if they find the work interesting. Yet, if people don't understand your "story" (the problem you're solving, its importance, and your approach to solving it), they will never make it past the introduction into the details of the paper. There are two schools of thought for writing the introduction: if you write it first, you ensure that you have a clear story. If you write it last, then your introduction will clearly reflect the ultimate story and be bolstered by your results.

My advice is to *do both*. Writing the introduction is probably the most difficult part of writing the paper, which may come as a surprise to new students. After all, the introduction is high-level and often doesn't reflect the nitty gritty details of how difficult that experiment was to run, and how much toil went into writing the system. To the contrary, *distilling your work* from those details so that the rest of the world can digest the takeaways—stepping out of the weeds and trying to figure out why your work matters—can take more effort than anything. The same results can be presented in many different ways, so you will likely need to perform many iterations on your story. Typically, you can "test" your story out on other colleagues and readers with a draft introduction in parallel with completing the rest of your work.

Write the introduction first—and last. I typically try (and encourage students) to write a draft of the paper's introduction *before* writing other parts of the paper. You will most likely eventually discard this draft completely, since it's very hard to predict what one should be emphasizing in the introduction before the "meat" of the paper is done. Yet, writing an early introduction draft allows a researcher to put the work in context, and to argue (1) what the problem is, (2) why it is hard, and (3) why the solution (if it is achieved) will be interesting to readers. If an introduction draft can't articulate answers to these questions, even at an early stage in the research, then the work likely needs some focus, as well. When you write the second version of your introduction (after the rest of the paper is written), you are in a good position to refine your story, *making sure that any claims you make are clearly supported by your results and data*. Your results should be stated as clearly and specifically as possible. You should neither understate your results (which might cause the reader to lose interest and stop reading) or overstate them (which might cause the reader to be annoyed or "let down" when they dig in and find that you don't deliver on what you've promised).

Perform some "landscaping" on your paper. Your paper should be pleasing to the eye. Care in preparation of the paper suggests that you also cared about the research itself. Conversely, if the preparation of the paper is sloppy, how can the reader trust that you were not similarly sloppy in writing code, setting up and running your experiments, and analyzing your results? Therefore, take some time to make your paper look nice. Here are a few quick tips on "landscaping" your paper.

- **Eliminate widows.** Section headings at the bottoms of columns or pages look sloppy. Depending on your word-processing tool, there are good ways to get rid of widows. In Latex, you can always use a "\vfill" to kick the section heading to the top of the next column. Similarly, widows on paragraphs (paragraphs that end with a single word on the last line) look sloppy. Typically you can eliminate such widows by wordsmithing (see below on omitting needless words). Latex also has macros such as "\widowpenalty" that help ensure that the text layout avoids widows whenever possible.
- **Place (and create) signposts, figures, and graphs to create whitespace and avoid "walls of text".** Reading long installments of technical writing is tiring, and seeing a page of uninterrupted technical writing can psychologically tire the reader before they even begin to read. Create whitespace and take other steps to break up large chunks of text. Use paragraph headings to break up long strings of text (ensuring that your headings match the topics you've outlined, of course). Try to space figures, tables, and graphs throughout the paper so that most pages have a mix of figures and text. If you don't have a figure to break up the text, consider making one. Often, a table summarizing results can be useful. Or, you might find that a figure (or block of pseudocode) describes a concept better than a paragraph in your writeup. Take steps to ensure that your writeup is not a long stream of

text.

- **Check your spelling.** Run a spell checker. Every word-processing tool has one, and a writeup that is rife with spelling errors reflects carelessness. Again, if you can't be bothered to check your spelling, why should a reader trust that you spent time debugging your experiment or code? While a reader can forgive grammatical errors, especially when the writeup comes from a non-native English speaker, anyone can run a spell checker, so spelling errors simply reflect carelessness. That said, not all spell-checkers are perfect, and sometimes typographical errors actually manifest as real words; spell checkers won't catch those, so they should be used only as a first pass.
- **Make the last page look decent.** If you've edited your text down to remove redundancy, you may find that you have some extra space at the end of your paper. Try to adjust your formatting so that the last page of the paper looks decent. If your paper is double-column and you end with references, balance the references across the two columns so that both columns are both equally (partially) filled. Latex has a "\balance" command that you can use to balance the columns on the last page of your writeup. If you have a bit more time, you can also make small adjustments to margins, space between columns, space around figures, sizes of figures, and so forth to ensure that the last page doesn't look like you just suddenly decided to stop writing.

Be Efficient

"I'm sorry I have had to write you such a long letter, but I did not have time to write you a short one." – Blaise Pascal

After you've made a good impression, you still need to engage the reader at least long enough to convey your main message. You should assume that the reader is busy, has a short attention span, and (unlike the situation when they are, say, reading a news article or novel) simply wants to learn what you've done as quickly as possible and then move on. In contrast to writing a novel, your goal is not to hold the reader in suspense or to illustrate your command of vocabulary or literary construction. Rather, the goal is to transfer information as efficiently as possible.

Tailor the length of your paper to the information content. People sometimes feel compelled to stuff a paper to the conference's or journal's page limit (after all, if the paper is too short, isn't there "more work to do"?). I've had a few papers that do fall short of the page limit, although I've found that this scenario is the exception rather than the rule (see Blaise Pascal: it takes time to shorten a paper!). Remember that your goal is to *efficiently transfer information*. Therefore, economize on words. Do not add unnecessary sentences, paragraphs, or other content that do not support a topic or point that you've explicitly outlined.

Keep words simple and sentences short. The easier it is to read your paper, the more likely a reader will keep reading. Therefore, don't make the reader work harder than they have to. On a sentence level, follow Strunk and White's most indispensable advice: *Omit needless words* (*e.g.*, "to" as opposed to "in order to", "optimizing" instead of "the problem of optimizing", "This module…" instead of "This is a module that", "more quickly" instead of "in a shorter

time", "whether" instead of "the question as to whether"). Avoid padding phrases such as "the fact that" and unnecessary adverbs (e.g., "eliminated" instead of "totally eliminated", "fast" instead of "very fast"). Keep words simple (e.g., "use" instead of "utilize", "new" instead of "novel"), and keep sentences short.

Eliminate redundancy. Always make at least one (and probably several) editing passes over your paper to cut out redundancy. You likely (and hopefully) did not write your paper in a linear fashion, and different co-authors may have written different sections or pieces of the paper. It's likely that you've stated the same point multiple times throughout the paper. Cutting redundancy takes practice, and it sometimes requires divorcing the editing process from your ego (after all, isn't that sentence you wrote beautiful? It'd be a shame to cut it, right?). If you need help cutting, ask someone else (a co-author or colleague) to help you—they're not as married to your text as you are.

Be as specific and precise as you can. Precision is especially important when describing results, and when writing up an experiment or evaluation section. For example, a good rule of thumb for writing up an evaluation section is: Would someone be able to reproduce your experiment from your description? If not, you need to be more clear. In an evaluation section, clearly state your assumptions; in what context(s) do your results hold? How general are they? Clearly describe your evaluation setup, as it sets the context for the results themselves. What machines did you use to run the evaluation? What data did you gather to support your conclusions? What algorithms did you use to analyze the data? *The evaluation setup should read like a recipe.* Someone reading it should be able to reproduce what you've done, at least well enough to understand the context of the results. Similarly, the results should be clearly and precisely described. Avoid writing sentences such as "This optimization significantly reduced the system's running time." Such vague statements frustrate the reader. Instead, substitute precise and specific information: "This optimization reduced the system's running time by 20%".

Use clear and consistent terminology. Unclear, vague, and inconsistent terminology is incredibly frustrating and also typically reflects muddy thinking. If you call something a "node" in one part of your paper, a "host" in another part of the paper, a "client" in another part, and a "machine" in yet another part, people will have absolutely no clue that all of these terms are in fact referring to the same thing. Before you get too far along in the writing process, clearly define your lexicon. Defining your terms up front can help clarify your own thinking, by helping you to identify and articulate the important concepts that you will present. Ultimately, this clarity will be reflected in your writeup.

Reduction to Practice: Write Every Day

Writing a technical paper involves clearly presenting a good story, as efficiently as possible. The above tips should help you compose your story and ensure that it "flows" and ultimately produce a final result that leaves a good impression and communicates your results as efficiently as possible. But, simply reading this post won't make you a good writer: putting these tips into practice takes tireless repetition. Try writing something every day to practice these tips. Keep

a "running draft" of your research paper. Start a blog and practice composing posts that present shorter and simpler ideas and thoughts than a full research paper. Keep a diary or notebook. Ultimately, becoming a good writer requires relentless, *daily* practice, with continual attention to the above tips.

One Comment on "Storytelling 101: Writing Tips for Academics"

1. *The Paper Reviewing Process | How to Do Great Research* says:
   October 18, 2013 at 6:32 am
   […] Storytelling 101: Writing Tips for Academics → […]

Blog at WordPress.com. The Clean Home Theme.