Natural patterns: How native and non-native speakers of English can avoid unnecessary complexity in scientific writing

by David Alexander

Synopsis

Scientific writing would be much clearer if it followed the patterns of ordinary written English. Using examples written by non-English-native researchers in biomedicine, I show that writers could build arguments more effectively by ensuring that the sentence topic is consistently the same as the sentence subject. This would also enable them to disregard much spurious advice on the active and passive voice. Writers should also express action using verbs, not nouns; and should inflict complexity on their readers only when it is unavoidable. These rules apply across all scientific disciplines.

Even in their own language, few people find writing easy. In the first session of my writing courses for PhD students in biomedicine, I am therefore unsurprised when most participants—all second-language writers—admit that they don't really enjoy it.

Neither do they always enjoy scientific *reading*. Many published texts yield their information unwillingly, often forcing their readers to read the same passage more than once—not because the science is difficult to understand, *but because the writing is*. As a result, scientific readers must quickly become skilled in the art of decoding—in working out, by intuition and experience, *what the writer probably meant to say*.

This may be a practical necessity, but it raises an important question: why should such information be encoded in the first place? Because this is certainly what happens: in my work I see it regularly in the writing of young researchers, who, often misguided by older colleagues, struggle to achieve properly 'scientific' formulations—by which I mean the empty and redundant constructions that clutter, confuse and deaden so many published articles.

Sentence topic? Sentence subject!

As such attempts can lead to a caricature of scientific style, they can often be instructive. Take this example from the results section of a biomedical paper:

> *For several hypermethylated genes we observed down-regulated gene expression.*

This short sentence contains two problems. The first is one of absurdity: the redundancy of the word 'for'. An ordinary reader would expect this word either to refer to duration, which it does not; or to mean 'on behalf of', which would be nonsense ('*On behalf of several hypermethylated genes, we observed down-regulated gene expression*…'?).

But the second problem is much more important, as it is so common in scientific writing: the failure to ensure that the sentence *subject* is the same as the sentence *topic*. In ordinary, non-scientific language, subject-topic agreement is standard practice: as a result, the reader easily understands what a sentence is about.

Is science so special that different rules must apply to its description? No. For despite appearances to the contrary, the statement 'we observed down-regulated gene expression' is not *about* 'we' the researchers, but about the expression of several hypermethylated genes. All the writer needed to do was to tell us what this expression *was* or what it *did.* Did she doubt her observations? No. There was therefore no need to resort to a pseudo-scientific ploy ('well, this is what I *think* I saw, but of course in science you can never be sure').

When asked, the writer was in fact relieved to be allowed to put things plainly, and confirmed that she had really wanted to state a simple fact:

> *The expression of several hypermethylated genes was down-regulated.*

Failure to follow the principle of subject-topic agreement is devastating for scientific writing. First, the rule is essential to argument-building. Second, as William S. Robinson has pointed out [1], its proper application invalidates a lot of vague and illogical advice on the use of the active and passive voices. The following example illustrates both of these points:

> **Background**: *Congenital heart defects (CHDs) have a multifactorial aetiology, in which subtle genetic factors and periconceptional exposures interact. In this aetiology, derangements in the homocysteine and detoxification pathways may be involved. The recently identified nicotamide N-methyl transferase (NNMT) gene and its substrate nicotinamide play a role in both pathways.*

This paragraph starts out well, but quickly loses focus: the second sentence does not seem to follow on from the first. Although this starts by alluding to 'old' information—the aetiology—an allusion is not explicit enough. The point is that this sentence is *about* the aetiology. However, this is not immediately apparent, because derangements—not the

>>>

aetiology—are the subject of the verb ('may be involved'). In other words, there is an important mismatch between grammatical subject and sentence topic.

The resulting dissonance sows seeds for confusion, which is compounded when the third sentence does not link up clearly with the second. Even though they have been slightly wrong-footed, readers might still expect the writer to develop the idea of pathways, which was introduced at the end of the previous sentence. Instead, the grammatical subject contains new information—NNMT and its substrate—for which they are entirely unprepared.

If the two mismatches between sentence topic and subject are corrected, the true topics are allowed to express themselves as they should:

> **Background**: *Congenital heart defects (CHDs) have a multifactorial aetiology, in which subtle genetic factors and periconceptional exposures interact. This aetiology may involve derangements in the homocysteine and detoxification pathways. An important role in both pathways is played by the recently identified nicotamide N-methyl transferase (NNMT) gene and its substrate, nicotinamide.*[1]

In this way, the paragraph gains its intended focus: the NNMT gene and its substrates—which were in fact nothing less than the theme of the ensuing article.

Active or passive? Let the topic decide

These changes bring me to my point about the active and passive voice, which is illustrated in the corrected version of the third sentence, '[a] role in these pathways is played…'.

Rightly, this sentence is now *about* its subject. So if writers are to observe the rule *that the main verb should always be the one linked to* the sentence subject, how much freedom do they have to choose between the passive and the active voice? None! This is because the voice is not actually chosen by the writer, but dictated by the sentence subject.

This invalidates widely published injunctions to writers on the use of the active and passive voice. It makes no more sense to 'always' use the passive than it does to 'mainly' use the active, or to 'judiciously balance' the active and passive. All such advice fails to take account of the fundamental role of voice in an English sentence. Voice selects itself according to what is being said! If this is not properly understood, a misguided 'principle'—or is it a prejudice?—will prevail over linguistic practice. And principles based on a fundamental misunderstanding of language provide a poor basis for good advice.

Nouns can't do the work of verbs

Scientific writing is badly affected by a third misunderstanding. The following sentence was presented in an abstract by one of our students, who added—in my view significantly—that no-one in her department had been able to get it 'right':

> *FK506 may be useful to facilitate retention of chondrocyte phenotype for cell-based therapies, or for in-vivo application to enhance the repair of focal chondral lesions and may contribute in treating osteoarthritis.*

Though the solution was actually extremely simple, simplicity does not present itself easily in disciplines where complex language is assumed to be *de rigueur*. When the student was asked 'What precisely is it that FK506 may *do*?', she was immediately able to reformulate the sentence. In the redrafted abstract, it read as follows:

> *FK506 may help maintain the chondrocyte phenotype, which is essential for cell-based therapies. In vivo, it may also improve the repair of focal chondral lesions.*

The clue here, of course, is not just her substitution of the simple verb construction 'may help maintain' for the much heavier 'to facilitate retention of': it is also her use of another simple verb construction—'is essential for'—which allows her to end the first sentence with a new and much more powerful idea. This sets the scene for a second, verb-based sentence ('it may also improve…'), which now lends much greater power to the second idea.

Such simple but effective corrections highlight the profound failure of so much scientific writing to capitalise on the power and dynamism of verbs. For, at best, most writers seem to use verbs grudgingly: as a necessary grammatical evil—an irrelevant, barely scientific element whose only function is to ensure that a sentence is nominally a sentence.

The result is the endless repetition of a limited range of empty, lifeless and *unnecessary* verbs—'showed', 'was found', 'were observed'. Whether the voice is passive ('measurements of bile accumulation were conducted in splenetic toads') or active ('we performed quantification of excessive nominalisation by science writers'), the result is not only turgid, but also vague. For if science is supposed to be all about precision, verbs can help you achieve it: 'heart rate was measured…' or 'we quantified…'.

Distinguishing between terminology and jargon

The fourth misunderstanding concerns an apparent inability or unwillingness to distinguish between scientific terminology, which is necessary and therefore unavoidable, and wordiness or jargon, which are not. By often expressing their dislike of what John Kirkman calls 'the unnecessary use of specialised terminology', my students are no differ-

[1] On several occasions, these two sentences have been shown to groups of PhD students from non-biomedical disciplines, who were therefore unfamiliar with the vocabulary. Though few understood the content, nearly all found the second version to be more accessible than the first. In short, good writing is not just about key words; it is also about how they are connected.

ent from the native-speaking members of the Biochemical Society who participated in Kirkman's test of style preferences [2].Why do scientist readers dislike jargon? Because it blurs, confuses or even hides the message.

Many writing textbooks warn against both jargon and wordiness. On occasion, so do the journals, though I wish that advice such as the following were published—and heeded—much more widely:

> Strive for clarity above all else. Avoid unnecessary jargon. If a $1 word will do the job, choose it over a $10 word. Readers will find a clear manuscript more persuasive and enjoyable than one that attempts to make its authors sound scholarly. [3]

With reference to wordiness, the same editorial—in *The Archives of Pediatrics and Adolescent Medicine*—continues:

> Brevity is a virtue. Omit unnecessary phrases, sentences, and paragraphs. Carefully examine your work for empty phrases. For example, 'a majority of' can be replaced with 'most.' 'It is of interest that' can just be deleted. Search for useless sentences at the start and end of paragraphs; they are often lurking there.

Writers should also be urged to apply Kenneth Hudson's 'key test':

> '[ask yourself]: "Could this have been expressed more simply without communication suffering in the process?" If the answer is "Yes" then you've got a piece of jargon.' [4]

Four reminders and a warning

Non-English-native-speaking scientists may be pardoned for assuming that scientific writing is an arcane science of its own. It is not. Most editors (for, unfortunately, there are exceptions) are like nearly everyone else: whatever their discipline, they seek brevity, transparency and clarity. By observing the principles outlined above, writers might go a long way towards achieving it.

Naturally, specialised vocabulary and formulations play vital roles in science writing—but the main story in all genres and disciplines *is nonetheless told by the underlying linguistic structures*. These principles should therefore apply across all academic disciplines. Used sensitively, they provide guidelines that are just as relevant to proteomics—for example—as they are to psychiatry.

For my students, who are indeed drawn from a very wide range of disciplines, I sometimes summarise these points as follows:

1. Remember that the first purpose of scientific language is to be clear, precise and accurate (it being understood that precision and accuracy are not the same thing).
2. The only difficult words in a scientific text should be the scientific ones, which were invented only because ordinary language has no other words to describe what needs to be described.
3. Any other words you use should be the same as those you'd use in the ordinary language. Few readers want texts to contain unnecessary words (or unnecessarily complicated words), as they make the text harder to understand. People who *do* want such words are simply being pretentious.
4. To the greatest possible extent, all these ordinary words should be combined with the scientific ones according to the patterns and constructions of ordinary language.

Ever the teacher, I also end with a warning:

Failure to follow points 1 to 4 has led to what people mistakenly call 'scientific' language. Much so-called 'scientific' language has therefore developed as a result of misunderstandings about the kind of style that is required. All too often, an unnecessarily complex approach leads to less clarity, precision and accuracy. Sometimes it makes scientific communication totally unreadable.

Acknowledgements

Mike Gould suggested I write this article. He and Joy Burrough-Boenisch commented on the manuscript and recommended a number of vital changes. Thanks are also due to Dominique Stumpel, Anna van der Windt and Lydi van Driel for their permission to use passages from early drafts of their writing; and to Lenie van Rossem for some very useful questions and a very useful reference.

David Alexander
Baarn, Netherlands
david@waywords.nl

David Alexander teaches biomedical writing at Erasmus University Medical Centre, Rotterdam, and is director of Ways with Words English Language Services, Baarn.

References:

1. William S. Robinson, *Sentence Focus, Cohesion, and the Active,* TETYC, May 2000
2. John Kirkman, *Good Style: Writing for Science and Technology*, Routledge, Abingdon, 2004 p. 193
3. Editorial (p. 1000) Arch Pediatr Adolesc Med/Vol 155, Oct 2001
4. Kenneth Hudson, *The jargon of the professions*, Macmillan, London, 1978.

Short guide to cancer symptoms and treatment

An article published by BBC News in conjunction with Cancer Research[1] starts with the alarming statement that one in three of us will be diagnosed with cancer during our life. While the number of new cases of lung cancer is actually falling following the trend of fewer smokers, cancer is becoming more common overall. The good news is that treatment is improving. The article gives a brief but useful guide to the symptoms of the different cancers and possibilities for treatment.

[1] http://news.bbc.co.uk/2/hi/health/3444635.stm