DOI:10.16262/j.cnki.1005-0841.2003.02.022

On Writing Scientific Articles in English

James Hartley
(Keele University, UK)

1 The Nature of Scientific Prose

If we examine the text of scientific articles it is clear that there is a generally accepted way of writing them. Scientific prose in English stereotypically uses the third person, the passive tense, complex terminology, and various footnoting and referencing systems. Scientific prose is not known for discursive anecdotes, humour, pictures, colour, bizarre typography or exclamation marks! Often the written text appears quite impersonal - the human element is removed[1].

These matters are important when it comes to teaching people how to write scientific articles. Consider, for example, the following advice from a text on writing for psychology students[2]:

"Good scientific writing is characterized by objectivity. This means that a paper must present a balanced discussion of a range of views.... Moreover, value judgements, which involve moral beliefs of what is 'right' or 'wrong' must be avoided... The use of personal pronouns is unnecessary, and can lead to biases or unsupported assumptions. In scientific papers, therefore, personal pronouns should not be used. When you write a paper, unless you attribute an opinion to someone else, it is understood to be your own. Phrases such as "in my opinion" or "I think," therefore, are superfluous and a waste of words... For the same reasons, the plural pronouns we and our are not used. " (pp.2-3)

In my view it is these very features of academic writing discussed above that (partly) make scientific text difficult to read. Making the text 'objective' and 'impersonal' makes it less accessible, and thus it hinders communication and understanding.

2 Measuring the Difficulty of Scientific Prose

One way of assessing whether or not this opinion is true is to compare the difficulty of scientific texts using computer-based readability programs. These programs, although somewhat crude, allow one to compare different versions of text in terms of how difficult they are to read. Table 1 provides the data that I obtained when I applied one such package (MicroSoft Office '97) (1)to the abstract of this article and (2) to the text of the whole of this section that you are reading now.

Table 1 Readability measures for extracts from different parts of the texts provided in this article

Measure	Current Article		Watson &
	Abstract	Extract	Cric(1953)
No. of words per sentence	25.0	22.3	18.2
Percent of passive sentences	100	5	12
Flesch Reading Ease Score	41.5	49.3	56.1

The abstract, in my view, is more typical of scientific writing than is the current text, and this is reflected in the Flesch Reading Ease scores. Table 1 also supplies similar data from the same measure applied to the opening paragraphs of Watson and Crick's famous paper[3] announcing their discovery of the structure of DNA (see Panel 1).

In all three cases we need to know that the Flesch scores range from 0 to 100, and that the higher the score the more readable the text. In point of fact, scores of 0—29 are labelled 'very difficult', and 30—49 'difficult' to read[4].

The data for Watson and Crick's paper show, nonetheless, that some scientific papers can be highly readable. Similarly, Einstein's first paper on relativity, published in 1905, illustrates the same point. References are eschewed, the style is simple, and the only acknowledgement is to a friend for being helpful in discussions leading to the writing of the paper[5]. Such papers are more direct and personal.

More detailed studies of less familiar papers show much the same thing. In one such enquiry, my colleagues and I used two computer-based measures to assess the readability of 42 highly influential and 42 less-influential research articles in psychology[6]. We

found that articles judged by colleagues to be influential had a higher Flesch score (average 33) than did control articles published in the same journals (average 25).

Data from less influential articles are rather more depressing. Table 2 shows the results from the same measure used in Table 1, this time applied to the Abstracts, and extracts from the Introductions and Discussions of 80 research articles published in the *Journal of Educational Psychology* [7]. It can be seen that, by and large, the Abstracts are the most difficult to read, and the Discussions are the easiest. But the average Flesch score for the Discussions is still only 23.

Table 2 Readability measures for the Abstracts, and for extracts from the Introductions and Discussions of 80 articles published in the Journal of Educational Psychology

Measure		Abstracts	Introductions	Discussions
No. of words	m	22.8	27.4	25.2
per sentence	s. d.	4.6	4.5	4.4
% of passive	m	28.8	17.5	19.1
sentences	s. d.	23.3	13.9	11.3
Flesch Reading	m	18.1	20.5	22.7
Ease score	s. d.	12.2	9.7	9.1

3 Improving Readability

It seems clear that if we want to make scientific texts more accessible then we need to use simpler and clearer language. Indeed, people have been saying this for a long time now. Almost 100 years ago the Fowlers wrote[8]:

"Anyone who wishes to become a good writer should endeavour, before he allows himself to be tempted by the more showy qualities, to be direct, brief, vigorous and lucid".(p.1)

Such things are easy to say but difficult to do, so I have provided some guidelines in Panel Two. Other, useful, additional full-length texts are available [e.g., 9—11] and there are two journals (with online contents) that might be helpful[12,13].

Unfortunately it is more difficult for non-native speakers of English to write lucid prose than it is for native ones[14,15]. Regrettably, methods of automatic translation have not yet progressed sufficiently for us to be able to turn scientific articles written in different languages into formal scientific English to a satisfactory standard. And further difficulties will ensue if the English required is more colloquial and idiomatic[16]. Automated grammar and style checkers may help but, in my experience, you need to already have a good knowledge of grammar and style before you can judge the validity of many of their suggestions. Thus machines and humans have to work in combination[17].

John Swales and Christine Feak have written two useful books for non-native speakers of English in this regard[18,19]. Teachers of English for non-native speakers may also find the following articles of interest[1,15,16,20] of interest.

However, in my view, non-native writers of English will best be aided by working with native speakers of English. Native speakers are more aware of the subtleties and nuances that might escape non-native English speaking colleagues. There is a case, therefore, for more international collaboration and assistance when authors with different nationalities are involved. Fortunately, such assistance is much easier today via e-mail and the Internet.

Acknowledgements I am grateful to many colleagues for their helpful comments on earlier versions of this article.

References

[1] Hyland K. Disciplinary discourses: writer stance in research articles. In C.N.Candlin and K. Hyland (eds), Writing Texts, Processes and Practices. London: Longman, 1999, 99—121.

[2] Smith T. R. Writing in Psychology. 2nd ed. New York: Wiley, 1996, 2-3.

[3] Watson J. D., Crick F. H. C. A structure for Deoxyribose Nucleic Acid. Nature, 1953, 171, April 25, 737—738.

[4] Flesch R. A new readability yardstick. Journal of Applied Psychology, 1948, 32: 221—223.

[5] Holton G. Thematic Origins of Scientific Thought. Harvard: Harvard University Press, 1988, 212—213.

[6] Hartley J., Sotto E., Pennebaker J. Style and substance in psychology: Are influential articles more readable than less influential ones? Social Studies of Science, 2002, 32(2): 321—334.

[7] Hartley J., Pennebaker J. W., Fox C. Abstracts, introductions and discussions: How far do they differ in style? Scientometrics, 2003, 57(3): 389—398.

[8] Fowler H. W., Fowler F. The King's English. Oxford: Clarendon Press, 1906.

[9] Day R. A. How to Write and Publish a Scientific Paper (5th edition): Westport, Connecticut: Greenwood, 1998.

[10] Turk C., Kirkman J. Effective Writing (2nd edition). London: E. & F.N. Spon, 1989.

[11] O'Connor M. Writing Successfully in Science. London: Harper Collins, 1991.
[12] English for Specific Purposes. Elsevier.
[13] Journal of English for Academic Purposes. Elsevier.
[14] Martin P. M. A genre analysis of English and Spanish research paper abstracts in experimental social sciences. English for Specific Purposes, 2003, 22: 25—43.
[15] Yang Y. Y., Swales J. M. Informal elements in English academic writing: threats or opportunities for advanced non-native speakers? In C. N. Candlin and K. Hyland (eds), Writing Texts, Processes and Practices. London: Longman, 1999, 145—167.
[16] Milton J. Lexical thickets and electronic gateways: making text accessible by novice writers. In C. N. Candlin and K. Hyland (eds), Writing Texts, Processes and Practices. London: Longman, 1999, 221—243.
[17] Hartley J. The role of colleagues and text-editing programs in improving text. IEEE Transactions on Professional Communication, 1984, P-C 27(1): 42—44.
[18] Swales J. M., Feak C. B. Academic Writing for Graduate Students: A Course for Non-native Speakers of English. Ann Arbor, Michigan: University of Michigan Press, 1994.
[19] Swales J. M., Feak C.B. English in Today's Research World: A Writing Guide. Ann Arbor, Michigan: University of Michigan Press, 2000.
[20] Parkinson J. Acquiring scientific literacy through content and genre: a theme-based language course for science students. English for Specific Purposes, 2000, 19(4): 369—387.

Panel 1. The opening paragraphs from Watson, J. D. and Crick, F. H. C. A structure for Deoxyribose Nucleic Acid. Nature, 1953, 171, April 25, 737 — 738, reproduced with permission.

We wish to suggest a structure for the salt of deoxyribose nucleic acids (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey. They kindly made their manuscript available to us in advance of publication. Their model consists of three inter-twined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

Panel 2: Guidelines for clear scientific writing

1. Keep in mind your readers-they may not be experts

Imagine that you are writing for a fellow colleague -or for one of your students -who is familiar with the conventions of your discipline, but who does not know your area. Readers need to be able to grasp what you did, what you found, and to follow your arguments easily.

2. Use the first rather than the third person

Compare: ''We suggest that...' with 'This paper suggests that...'

3. Use short, simple words

It is easier to understand short, familiar words than technical terms which mean the same thing. Compare: `We assume, from the start...` with: `We assume, a priori...`

4. Use active tenses

It is easier to understand text when writers use active tenses rather than passive ones. Compare: 'We found that the chemists varied more than the engineers on a measure of extraversion...' with ''Greater variation was found on a measure of extraversion with the chemists than with the engineers...'

5. Sequencing in sentences

It is more helpful for the reader in English if the subject of the verb comes before, and the object after, the verb. Compare: 'Students need accessible information to become intelligent customers of an increasingly diverse provision and to meet their own diverse needs...' with: 'To become intelligent customers of an increasingly diverse provision, and to meet their own diverse needs, students need accessible information...'

6. Avoid negatives

Negatives, especially double or treble ones, can be confusing. Compare: 'The figures provide no indication that the costs would not have been lower if competition had not been restricted...' with: ''The figures provide no indication that competition would have produced higher costs...'. Negative qualifications can be used, however, for particular emphasis, and for correcting misconceptions.

7. Avoid abbreviations

Many writers use abbreviations for technical terms: e.

g., RAE for Research Assessment Exercise. Too many abbreviations are off-putting. Furthermore, if the abbreviations are unfamiliar, it is easy to forget what they stand for.

8. Avoid overloading the text with references

It is difficult to read sentences that end with long lists of supporting references. It is better to cite only the more recent papers that between them summarise earlier research. Compare: 'Common practice has been to assume the condition of local equilibrium (for example, see Bickle and others, 1997, and Brady, 2001, for surveys of this research)...' with: 'Common practice has been to assume the condition of local equilibrium (Baumgartner and Rumble, 1988; Bickle and Baker, 1990; Bickle and others, 1995, 1997; Brady, 2001, Cartwright and Valley, 1991; Ferry, 1986, 1994)...'

9. Vary sentence lengths

It is easier to understand short sentences than it is to understand long ones because long sentences overload the memory system. Short sentences do not. However, it is good practice to vary sentence lengths, as long strings of short sentences feel 'choppy'.

As a rule of thumb I suggest that sentences less than 20 words long are probably fine. Sentences 20–30 words long are probably satisfactory. Sentences 30–40 words long are suspect. Sentences with over 40 words in them will probably benefit from re-writing.

10. Use short paragraphs

Short paragraphs are easier to read than long ones. Any typescript that has a page of text without at least one new paragraph needs attention!

11. Place sequences in order

It is best to describe procedures in the order that they take place. For example, compare: 'Before the experiment commenced, we first briefed the participants on the necessary procedures and any precautions that they should take...', with: 'We briefed the participants on the necessary procedures and any precautions that they should take before the experiment began...'

12. Use numbers or bullets

Numbers or 'bullets' are useful if you want to make a series of points within a paragraph. Compare: 'Four devices to help the reader of a thesis are a detailed contents page, skeleton outlines for each chapter, headings in the text, and a concluding summary... with:

'Four devices to help the reader of a thesis are: (1) a detailed contents page; (2) skeleton outlines for each chapter; (3) headings in the text; (4) a concluding summary...'

It is best to use numbers when there is an order, or sequence in the points made, and bullets when each point is of equal value.

13. Use headings

Headings label sections so that writers and readers know where they are, and where they are going. Headings help the reader to scan, select and retrieve material, as well as to recall it.

Headings can be written in the form of statements or in the form of questions. If the headings are in the form of questions then the text below must answer them. This helps the author to present – and the reader to follow – the argument.

14. Print out and revise/edit draft copies

Print out draft copies when the text is nearing completion. Copies allow you to check more easily the tiny details–punctuation, references, etc., –as well as to get a better feel for the document as a whole. Think about global revisions – re-sequencing major portions – and local revisions–making changes to individual words and sentences.

15. When in difficulty...

If it is difficult to explain something, think of how you would explain it to someone else. Think of what you would say, try saying it, and then write this down. Then polish it.

16. Read the text out aloud

Reading the text out loud (or silently) to oneself is a useful way of seeing how well the text flows. You may find that you need to insert commas to make text groupings clearer, you may get out of breath because sentences are too long, and you might inadvertently read out a simpler version of the written text. If you do this, change the text to this simpler version.

17. Ask other people to read your drafts

Colleagues and students may be willing to read and comment on drafts. Ask them to point out those sentences or sections that they think that other readers might find it difficult to follow. People are more willing to point out difficulties for others than they are to admit to having difficulties themselves.

18. Revise continuously...

Never regard the latest version of the text as the final one. Put this version on one side and then come back to it a day or two later. Seeing the text with fresh eyes somehow suggests further changes. But draw the line eventually!