Does length matter?

The Word-Count Myth

Those who come to technical writing from a scientific or technical background are often puzzled by the ready acceptance of guidelines that have never been subjected to critical scrutiny. They read that the active voice is always to be preferred, that paragraphs must be kept short, that numbers must be spelt out if less than ten, and so on—and they search in vain for the research that might give these guidelines the imprimatur of scientific rigour. They find plenty of language pundits eager to share their beliefs, but far fewer experts offering sound reasons.

One set of guidelines crying out for scrutiny has to do with sentence length. Many a pundit is convinced that the number of words in a sentence should fall within a particular range. Some also contend that sentences should not have more than a specified number of words. For example, Dr Anetta Cheek tells us that we should "aim for an average sentence length of between about 15 and 22 words". Moreover, we should "avoid sentences of more than 40 words".[1] Dr Neil James, Director of Plain English Foundation, is also convinced that average sentential word-count is an important consideration in good writing: "a 15–20 word average is *fundamental* to writing well".[2] Further, James hints that 35 words is probably the maximum we should include in a sentence: "if your content is complex, or if you need to include a subsidiary point, you will want to go over 18 words—perhaps even stretching towards 35 words".

The view that sentential word-count is an important consideration in writing well has such a long history that it is surprising that it has received little critical attention. The view does have prima facie plausibility. We do get lost in long sentences (and probably no-one has ever fully understood that famous 823-word sentence in Victor Hugo's *Les misérables*). We attempt to make sense of a sentence with the mechanisms of our short-term (or working) memory. That memory is limited by our biology. If a sentence we are reading surpasses that limit, ideas we read earlier in it are pushed out of memory and we forget part of what we have just read. We will need to process the sentence in parts and this will require rereading (perhaps a number of times). The respectful writer—keen to communicate while imposing the least effort on readers—will try to avoid making readers reread, and they will do so by ensuring that the sentences they write do not exceed the capacity of our short-term memory.

The common view is that we do this by limiting our word count. Yet experiments that challenge this view are simple to construct. In my writing classes over the last few years I have displayed on a screen a simple 11-word sentence. The sentence remained visible for six seconds (more than enough time for the sentence to be read by any competent reader). Students were asked to read the sentence just once and then to write down its main points. They were told not to fuss about reproducing the sentence exactly as it was presented, but just to jot down its gist. I then displayed another 11-word sentence for six seconds. Again students were asked to jot down the main points. Here are the two sentences:

> [A] The United States of America has reluctantly signed the peace accord.
>
> [B] The hot, treeless plain is covered in dark, smooth, elongated rocks.

I then gathered the cards the students had written on and tallied the number of points that had been correctly recalled. Despite the sentences being of identical word-count and of comparable vocabulary, average recall scores—that is, the average number of correct responses given—were markedly

---

[1] A Cheek, "Defining plain language", *Clarity: Journal of the international association promoting plain legal language*, no. 64, November 2010, p. 10.

[2] N James, *Writing at work*, Allen & Unwin, Crows Nest, 2007, p. 244. Emphasis added.

different: sentence [A] a little over 89%; sentence [B] just 33.9% (from 519 students). If word count is a critical factor in reader comprehension, then surely these sentences should have yielded similar recall scores.[3] The explanation for the discrepancy is simple, and it does have to do with the limits of our working memory.

The notion of a limit to working memory was made popular in a 1956 literature review by the American psychologist George Miller. Miller thought the limit, averaged over diverse stimuli (such as words, numbers, sounds, etc.), was *seven* distinct chunks. Subsequent research indicates that the average is closer to four (a fact that torpedoes a major claim of the Information Mapping fraternity):

> "the capacity of short-term memory … is known to be quite small, only about four chunks."[4]
>
> "a central working memory faculty is limited to 3 to 5 chunks for adults."[5]

To psychologists, a *chunk* in the act of verbal comprehension is a basic unit of meaning. Moreover:

> "the basic units of meaning are propositions. Propositions are n-tuples of *word concepts*, one of which serves as a *predicator*, and the remaining ones as *arguments*, each filling a unique semantic role. The predicator specifies a relationship among the arguments of a proposition. For instance, in the proposition (LOVE, Experiencer: GREEK, Object: ART) there are two arguments, GREEK and ART, and the predicator LOVE; in English this proposition could be realised with the sentence *The Greeks loved art*. It is important to note that the arguments of a proposition are *concepts* rather than words."[6]

In this example, the basic unit of information has the S–V–O structure, that is, subject–verb–object: The Greeks (S) loved (V) art (O). But there are other equally basic structures. For example, S–V (as in "Emily laughed"), A–N ("bacterial infection") and A–V ("loudly abused").

What cognitive science tells us, then, is that if we want our readers to understand our sentences on one reading, we need to limit the basic units of information in them to no more than four. And this explains the result of the experiment I described above. To see how, dissect each sentence into its constituent chunks:

[A] The United States of America has reluctantly signed the peace accord.

Two chunks: <USP> and <SR> — that is, (a) the United States of America has signed the peace accord and (b) the signing was done reluctantly.

[B] The hot, treeless plain is covered in dark, smooth, elongated rocks.

Six chunks: <P, C, R>, <P, H>, <P, T>, <R, D>, <R, S>, <R, E> — that is, (a) the plain is covered in rocks, (b) the plain is hot, (c) the plain is treeless, (d) the rocks are dark, (e) the rocks are smooth and (f) the rocks are elongated.

The variance in the rates of recall—89% for [A] and 33.9% for [B]—is due to the fact that the sentences, though of identical word count, have significantly different chunk counts. Sentence [A] poses little cognitive strain on our working memory because it has just two chunks, well within the four-chunk limit of our working memory. On the other hand, sentence [B] forces the reader to attempt to squeeze six chunks into a four-chunk slot and, as a result, some chunks escape. In the experiment,

---

[3] Note too that the sentences used in the experiment have a word count well below what many language pundits—Cheek, James and others—consider the upper limit for immediate comprehension.

[4] W Kintsch & K Rawson K, "Comprehension", in M J Snowling & C Hulme, *The science of reading: A handbook*, Blackwell, Oxford, 2005, p. 224

[5] N Cowan, "The magical mystery four: How is working memory capacity limited, and why?", *Current Directions in Psychological Science*, 2010, vol. 19, iss. 1, pp. 52

[6] W. Kintsch W et al., "Comprehension and recall of text as a function of content variables", *Journal of Verbal Behavior and Verbal Learning*, 1975, vol. 14, iss. 2, p. 196

78% of students correctly recorded both chunks of information in sentence [A] whereas only 1.9% correctly recorded all six chunks in sentence [B].

A fixation on word-count—average or maximum—is thus misleading. I could write a document where the average sentence-length is between 15 and 20 words—a so-called "fundamental" average—no sentence is longer than 35 words, the vocabulary is familiar (as is the grammar and punctuation) and yet every reader will have to read every sentence at least twice to fully understand it. All I need do is pack more than four chunks of information into every sentence. My writing meets the pundits' guidelines, but it could hardly be considered good writing.

So length does matter—*so long as we are measuring the right thing*. And the right thing is not words, but chunks: the basic units of information that one is trying to impart to readers.

As we've seen, short sentences can be unreasonably burdensome on readers; conversely long sentences—longer than the limit recommended by many pundits—can be a breeze, so long as the chunk limit is kept to four or less. Here is an example:

> The Democratic People's Republic of Korea is smaller than the People's Democratic Republic of Lao, the People's Democratic Republic of Lao is smaller than the Republic of South Africa, the Republic of South Africa is smaller than the Democratic Republic of the Congo and the Democratic Republic of the Congo is smaller than the United States of America.

The logical structure of this 58-word sentence is very simple: A < B, B < C, C < D and D < E, and its apparent wordiness does not impede its immediate comprehension. This is because the concepts are simple—mere countries—despite their names being multi-word compound nouns. It is pertinent here to repeat the last sentence in the quote above from Kintsch:

> "It is important to note that the arguments of a proposition [that is, a basic unit of information] are *concepts* rather than words."

Finally, consider the readability scores in Microsoft Word (scores that have been used by courts of law in the USA to justify awarding damages against writers). The Flesch reading ease score—and its derivative, the Flesch–Kincaid Grade Level—are based solely on the average number of words per sentence and the average number of syllables per word in a given text. So what Flesch scores do our two eleven-word sentences get? Sentence [A], with just two chunks of information, gets a score of 41.8 while sentence [B], with six chunks, gets a score of 72.6. So the sentence that the vast majority of readers will have to read twice before they fully understand it is deemed more readable than the sentence that nearly everyone understands immediately on first reading! Proof yet again that the readability formulas in Word are best thrown onto the scrapheap of pseudoscience.

As the technical writing profession lurches towards accreditation, it is vital that the knowledge on which writers will be judged is evidence-based, that is, stamped with the imprimatur of scientific rigour. It is vital too that those who contribute to standards on technical writing ensure that their recommendations shine with the burnish of proof. It will backfire on the profession if later we find that our cherished guidelines were at best hunches, old wives' tales or stylistic prejudices.

Geoffrey Marnell