

# Rules for writing scientific texts in English

Reto Spöhel, rspoehel@inf.ethz.ch

November 24, 2012

## Abstract

This document points out a few simple but important rules that you should observe when writing scientific texts in English. Some of them are pretty universal, while others were written with a German speaker in mind specifically.

*Feedback, corrections and additions welcome!*

## Contents

1	Basic writing rules	1
2	Being nice to the reader	2
3	A few important grammar rules	3
4	Typography	5
5	Things you (usually) don't say in English – and what to say instead	6

## 1 Basic writing rules

The following rules need little further explanation; they are best understood by looking at the example in the booklet by Knuth et al., §2–§3.

### Rule 1.1: Write texts, not chains of formulas.

More specifically, write full sentences that are logically interconnected by phrases like ‘Therefore’, ‘However’, ‘On the other hand’, etc. where appropriate.

### Rule 1.2: Displayed formulas should be embedded in your text and punctuated with it.

In other words, your writing should not be divided into ‘text parts’ and ‘formula parts’; instead the formulas should be tied together by your prose such that there is

a natural flow to your writing.

## 2 Being nice to the reader

Try to write your text in such a way that a reader enjoys reading it. That's of course a lofty goal, but nevertheless one you should aspire to!

### **Rule 2.1: Be nice to the reader.**

Give some intuition or easy example for definitions and theorems which might be hard to digest. Remind the reader of notations you introduced many pages ago – chances are he has forgotten them. Illustrate your writing with diagrams and pictures where this helps the reader. Etc.

### **Rule 2.2: Organize your writing.**

Think carefully about how you subdivide your thesis into chapters, sections, and possibly subsections. Give overviews at the beginning of your thesis and of each chapter, so the reader knows what to expect. In proofs, outline the main ideas before going into technical details. Give the reader the opportunity to 'catch up with you' by summing up your findings periodically.

*Useful phrases:* 'So far we have shown that ...', 'It remains to show that ...', 'Recall that we want to prove inequality (7), as this will allow us to deduce that ...', 'Thus we can conclude that ... . Next, we would like to find out whether ...', etc.

### **Rule 2.3: Don't say the same thing twice without telling the reader that you are saying it twice.**

Repetition of key ideas is important and helpful. However, if you present the same idea, definition or observation twice (in the same or different words) without telling the reader, he will be looking for something new where there is nothing new.

*Useful phrases:* 'Recall that [we have seen in Chapter 5 that] ...', 'As argued before / in the proof of Lemma 3, ...', 'As mentioned in the introduction, ...', 'In other words, ...', etc.

### **Rule 2.4: Don't make statements that you will justify later without telling the reader that you will justify them later.**

This rule also applies when the justification is coming right in the next sentence! The reason for this rule should be clear: if you violate it, the reader will lose valuable time trying to figure out on his own what you were going to explain to him anyway.

*Useful phrases:* 'Next we argue that ...', 'As we shall see, ...', 'We will see in the next section that ...', etc.

### 3 A few important grammar rules

**Rule 3.1:** There is (almost) *never* a comma before ‘that’.

It’s really that simple. Examples:

We assume that ...  
*Wir nehmen an, dass ...*

It follows that ...  
*Daraus folgt, dass ...*

‘thrice’ is a word that is seldom used.  
*‘thrice’ ist ein Wort, das selten verwendet wird.*

Exceptions to this rule are rare und usually pretty obvious. For example, you may end up with a comma before ‘that’ because ‘i.e.’ is spelled out as ‘that is’:

For  $p(n) = \log n/n$  we have ... However, if we choose  $p$  a little bit higher, that is  $p(n) = (1 + \varepsilon) \log n/n$  for some  $\varepsilon > 0$ , we obtain ...

Or you may get a comma before ‘that’ because there is some additional information inserted in the middle of your sentence:

Thus we found a number, namely  $n_0$ , that satisfies equation (13).

If the additional information is left out, the sentence has no comma:

Thus we found a number that satisfies equation (13).

(For ‘that’ as a relative pronoun, see also Rules 3.3 and 3.4 below.)

**Rule 3.2:** There is usually no comma before ‘if’.

Example:

A graph is not 3-colorable if it contains a 4-clique.  
*Ein Graph ist nicht 3-färbbar, wenn er eine 4-Clique enthält.*

However, if the ‘if’ clause comes first, it is usually separated from the main clause by a comma:

If a graph contains a 4-clique, it is not 3-colorable .  
*Wenn ein Graph eine 4-Clique enthält, ist er nicht 3-färbbar.*

There are more exceptions to these rules than to Rule 3.1, which is why we are not discussing them here. Just keep in mind: don’t put a comma before ‘if’ without good reason.

### Rule 3.3: Non-defining relative clauses have commas.

and

### Rule 3.4: Defining relative clauses have no commas.

In English, it is very important to distinguish between two types of relative clauses: defining and non-defining ones. This is a distinction you absolutely need to understand to write scientific texts, because mistakes in this area actually distort the meaning of your text!

It's probably easier to explain first what a *non-defining* relative clause is. A non-defining relative clause simply gives additional information *that could also be left out* (or given in a separate sentence). For example, the sentence

The WEIRDSORT algorithm, which was found by the famous mathematician John Doe, is theoretically best possible but difficult to implement in practice.

would be fully understandable if the relative clause were left out completely. It could also be rephrased as two separate sentences:

The WEIRDSORT algorithm is theoretically best possible but difficult to implement in practice. [By the way,] WEIRDSORT was found by the famous mathematician John Doe.

This is what a non-defining relative clause is. *Non-defining relative clauses are always written with commas.* As a corollary it follows that you cannot use 'that' in non-defining relative clauses (see Rule 3.1!). It would be wrong to write

~~The WEIRDSORT algorithm, that was found by the famous mathematician John Doe, is theoretically best possible but difficult to implement in practice.~~

A special case that warrants its own example is when 'which' is referring to the entire preceding sentence:

Thus inequality (7) is true, which implies that the Riemann hypothesis holds.

As before, this is a non-defining relative sentence (it could be left out) and therefore needs a comma.

So let's discuss *defining* relative clauses next. A defining relative clause tells the reader *which specific item the main clause is talking about*. Leaving it out either changes the meaning of the sentence or renders it incomprehensible altogether. Consider the following example:

The WEIRDSORT algorithm is difficult to implement in practice. In contrast, the algorithm that we suggest is very simple.

Here the relative clause 'that we suggest' cannot be left out – the remaining sentence would make no sense since the reader would not know which algorithm it is talking about. This is what a defining relative clause is. *Defining relative clauses are never written with commas.* Usually, you can use both 'that' and 'which' in defining relative clauses, although in many cases 'that' sounds better.

As a final example, consider the following sentence:

For the elements in  $\mathcal{B}$  which satisfy property (A), we know that equation (37) holds.

This sentence does not make a statement about all elements in  $\mathcal{B}$ , only about those satisfying property (A). The relative clause is *defining*. (Thus we could also use ‘that’ in place of ‘which’.)

In contrast, if we add a comma the sentence reads

For the elements in  $\mathcal{B}$ , which satisfy property (A), we know that equation (37) holds.

Now the relative clause is *non-defining* – it just mentions in passing that all elements in  $\mathcal{B}$  satisfy property (A). The main clause states that equation (37) holds for *all* elements in  $\mathcal{B}$ . See the difference?

## 4 Typography

**Rule 4.1: Use opening (‘) and closing (’) quotation marks correctly.**

In  $\text{\LaTeX}$ , the closing quotation mark is typed like a normal apostroph, while the opening quotation mark is typed using the French *accent grave* on your keyboard (the *accent grave* is the one going down, as in *frère*).

**Rule 4.2: Use hyphens (-) and dashes (—) correctly.**

A hyphen is only used in words like ‘well-known’, ‘3-colorable’ etc., or to separate words that continue in the next line (which is known as hyphenating).

The equivalent of the German *Gedankenstrich* is the dash, which is typed as two subsequent hyphens in  $\text{\LaTeX}$ . For example, in the title of Section 5, it would be wrong to use a hyphen instead of the dash. (Some English authors use the even longer emdash (—) instead, which is typed as three subsequent hyphens in  $\text{\LaTeX}$ . Usually this emdash is used without spaces around it—like so.)

## 5 Things you (usually) don't say in English – and what to say instead

*The following entries should not be taken as gospel – they don't necessarily mean that a given word or formulation is wrong under all circumstances (obviously, this depends a lot on the context). However, in nine out of ten instances the suggested alternative is the better word to use.*

<del>It holds (that) ...</del>	We have ...	<i>Es gilt ...</i>
(‘Equation (5) holds.’ is fine, though.)		
<del><math>x</math> fulfills property <math>\mathcal{P}</math>.</del>	$x$ satisfies property $\mathcal{P}$ .	<i><math>x</math> erfüllt Eigenschaft <math>\mathcal{P}</math>.</i>
<del>in average</del>	on average	<i>im Durchschnitt</i>
<del>estimation</del>	estimate	<i>Abschätzung</i>
<del>composed number</del>	composite number	<i>zusammengesetzte Zahl</i>
<del>with the help of</del>	using	<i>mit Hilfe von</i>
<del>surely</del>	clearly	<i>sicher, bestimmt</i>
<del>monotonously increasing</del>	monotonically/monotone incr.	<i>monoton steigend</i>
(Actually, in most cases ‘increasing’ is just fine.)		