

# Grammar and style notes for scientific writing

Wilhelmiina Hämäläinen

## 1 Scientific writing style

Main goal: exact, clear, and compact.
---------------------------------------

- Compact is usually clear!
- Other desirable properties: smooth and objective

### 1.1 Exact

- Word choice: make certain that every word means exactly what you want to express. Choose synonyms with care. Be not afraid of repetition.
- Avoid vague expressions which are typical for the spoken language. E.g. the interpretation of words which approximate quantities ("quite large", "practically all", "very few") depends on the reader and the context. Avoid them especially if you describe empirical observations.
- Make clear what the pronouns refer to. The reader shouldn't have to search the previous text to determine their meaning. Simple pronouns like this, that, these, those are often the most problematic, especially when they refer to the previous sentence. Hint: mention the noun, e.g. "this test".  
→ See Section Pronouns.
- Avoid ambiguous and illogical comparisons. These are often due to missing words or nonparallel structures. E.g. "Female students draw concept maps more often than male students."  
"The students' points were lower than the average computer science students."

→ See Section Parallel constructions.

- Antropomorfism: do not attribute human characteristics to machines or other inanimate things. E.g. a computer cannot understand data, an experiment cannot control variables or interpret findings, a table or a figure cannot compare results.
- Incorrect grammar and careless sentence structures can create ambiguities!

## 1.2 Clear

- Use illustrative titles which describe the essential in a chapter or a section.
- Write a brief introductory paragraph in the beginning of each chapter or section with subsections.
- Divide the text logically into sentences and paragraphs.
  - Direct, declarative sentences with simple, common words are usually best.
  - Paragraphs should be logically uniform and continuous.
 → See Section Sentences
- Place the adjective or the adverb as close as possible to the word it modifies.
  - See Sections Adverbs and Word order.
- Avoid **scientific jargon** = continuous use of technical vocabulary when it is not relevant.
- Write numbers as digits when they refer to sizes or exact measurements. Otherwise the general rule is to write numbers < 10 as words. Express decimal numbers with a suitable precision. See APA pp. 122-129.
- Use punctuation to support meaning.
  - See Section Punctuation and APA pp. 78-88

## 1.3 Compact

- Say only what needs to be said!
- Short words and short sentences are always easier to comprehend

- Weed out too detailed descriptions. E.g. when you describe previous work, avoid unnecessary details. Give a reference to a general survey or a review if available.
- Don't describe irrelevant or trivial observations (i.e. don't mention obvious things)
- Avoid wordiness, e.g.

"based on the fact that" → "because"

"at the present time" → "now"

"for the purpose of" → "for/to sg."

Notice: "reason" and "because" have the same meaning → don't use together!

- Use no more words than are necessary. Redundant words and phrases (which have no new information) should be omitted.
- Avoid too long sentences and paragraphs

## 1.4 Smooth

- Verbs: Stay within the chosen tense! No unnecessary shifts in verb tense within
  - the same paragraph
  - in adjacent paragraphs

→ See Section Verbs.

- Use verbs rather than their noun equivalents
- Prefer active to passive voice
- Avoid long noun strings!

Hint: sometimes you can move the last word to the beginning and fill in with verbs and prepositions

- Each pronoun should agree with the referant in number and gender.
- Transitional words help to maintain the flow of thought

- time links: then, next, after, while, since
  - cause-effect links: therefore, consequently, as a result
  - addition links: in addition, moreover, furthermore, similarly
  - contrast links: but, however, although, whereas
- Notice: some transitional words (while, since) can be used in several meanings → limit their use to their temporal meaning! (Use "because" instead of "since"; "although", "whereas" or "but" instead of "while", when there is no time connection.)
  - Use abbreviations sparingly, especially the abbreviations which you define yourself for technical terms.
    - See Section Abbreviations.
  - Do not use emphasis (italics) when it is not needed. Use syntax to provide emphasis.
  - Metaphors can sometimes help to simplify complex ideas. However,
    - Don't overuse them
    - Don't mix several metaphors in one sentence
    - Avoid cliches

## 1.5 Objective

- Use the 3rd person rather than the 1st person.
  - Use emotionally neutral expressions, e.g. "Students suffering from dyslexia" → "students who have dyslexia"
  - Use words which are free from bias (implied or irrelevant evaluation) Especially, be careful when you talk about
    - gender
    - marital status
    - racial or ethnical groups
    - disability
    - age
- See Section Gender-neutral language.

### Hints:

- Select an appropriate degree of specificity. When in doubt, prefer the more specific expression. E.g.
  - Instead of "man" use "men and women" or "women and men" to refer to all human beings
  - Instead of "old people" define the age group "ages 65-83"
  - Instead of "Asian" mention the nationality "Chinese"
- Differences should be mentioned only when relevant. Careless use of biased words can create ambiguities.

E.g. avoid the use of "man" as a generic noun or an ending for an occupational title. Otherwise it can imply incorrectly that all people in the group are male.

## 2 Verbs

Remember two important rules when you use verbs:

1. The number of subject determines the number of verb
2. Do not mix inconsistent tenses

### 2.1 Number and person

- When the subject is singular third person (she/he/it), the verb needs suffix -s (in the present, positive sentence). The auxiliary verbs have their own special forms (is, can, has, does).
- Be careful with special phrases:
  - "A number of new experiments were done" (plural)
  - "Plenty of time was spent..." (singular)
  - "A few data points belong to cluster X" (plural)
- Notice: when the subject is composed of a singular and a plural noun by "or" or "nor", the verb agrees with the noun that is closer.
- If the number of the subject changes, retain the verb in each clause. E.g. "The positions in a sequence were changed and the test rerun" → "The positions in the sequence were changed, and the test was rerun."

## 2.2 Tenses (temporal forms)

- Default: the present
- Past or present perfect (but not both) when you describe previous research (literature review)
- Past tense to describe the experiments and their results

- In scientific writing, the default is present (is). With present, you can combine perfect (has been) (and future, will be) if needed, but not the other tenses.
- Use past tense (was) only for good reasons. It expresses that something belongs to the past and has already finished. E.g. when you report your experiments.
- Past perfect (had been) is seldom needed. It is used, when you describe something in the past tense, and you refer to something which has happened before it. E.g.  
"We tested the system with data which had been collected in *Programming 1* course."
- Notice: Use "would" with care! It expresses a conditional action. E.g. "it would appear" → "it appears".

## 2.3 Active or passive voice, which person?

### 2.3.1 Use of passive voice

- In active voice the actor is known, while in passive voice it is unknown.
- In the basic form of passive ("sg is done"), you can express also the actor ("sg is done by sy"). Expressing the actor is always more informative!
- It is often recommended to prefer active voice, but in scientific writing passive voice is sometimes convenient. It allows us to draw the reader's attention to the phenomenon or the event, instead of the actor. E.g. "The probabilities are updated by Bayes rule", "The values are recorded every minute".

- Often the purpose determines the voice. Usually we want to begin with a familiar word and put the new information in the end. E.g. before an equation or a definition, we can say "The model is defined as follows."
- However, do not overuse passive, and do not chain passive expressions. As a rule of thumb, use only one passive per sentence
- Read Section 11 in Strunk: "Elements of style"! (link in the course page)

### 2.3.2 "It is" and "There is/are"

- A formal subject "it" is sometimes used in passive expressions: "It is often recommended [reference] that..."
- Typical verbs in this expression are: say, suppose, consider, expect.
- "There is/there are" is a similar expression, but now we don't need the passive. This expression is used when the real subject (what is somewhere) comes later and we haven't mentioned it before.

E.g. "There was only one outlier in the data set 1" v.s. "The outlier was in the data set 1."

- The verb is nearly always "be" (sometimes "exist" or something else)
- Notice that the verb follows the real subject's number.  
E.g. "There were a lot of outliers in the data set 1."
- "There is" expression is seldom needed in scientific writing, and often you can circumvent it:  
"The data set 1 contained a lot of outliers."

### 2.3.3 Other passive expressions

- "We" can be used as passive. E.g. "In Chapter X, we define the basic concepts." However, it is better to say "The basic concepts are defined in Chapter X."
- "You" is sometimes used as passive, especially in manuals. Don't use it in scientific text!
- "People" when you refer generally to people. Quite a vague expression, not recommendable!

### 2.3.4 Person?

- Basic rule: avoid the first person (no opinions, but facts). However, sometimes we can use "we" as a passive expression. Problem: whom you are referring to, if you write alone?
- Referring to yourself: you can talk about "the author". E.g. "All programs have been implemented by the author." Notice that I don't guarantee that your supervisor likes this! Some supervisors prefer "I".
- Gender-neutral language: when you refer to an unknown user, student, etc. try to use gender-neutral language.
  - The most common way is to say "she/he" or "he or she". Some authors are careful about the order of her/him, as well! E.g. you can use every second time "she or he" and every second time "he or she". Remember to put the other pronouns in the same order ("She/he tries her/his best")
  - "One" is neutral, but sounds often awkward. "The learner can define one's own learning goals"
  - Sometimes you can avoid the problem by using plural.

## 2.4 Other notes

- Do not use short forms "isn't, can't, doesn't", but "is it, cannot, does not".
- "be verb+ing" form when something is currently happening or takes some time. E.g. "Thread 2 can be started in the same time when thread 1 is still running"
- Some verbs require that the following verb is in -ing form:

{enjoy, avoid, succeed in, finish, keep, mind, practice, risk} + verb + ing
--

E.g. "Students enjoyed learning new things"

- Special phrases: "be used to", "be (un)likely to"

## 2.5 Noun syndrom

”Noun syndrom” = use of common verbs {be, do, have, make, ...} + a noun

E.g. ”We can get better understanding...”, ”Different people have different responses to the methods”

⇒ Prefer illustrative verbs!

**Task:** How would you correct the previous sentences?

### Useful verbs:

represent, analyze, compare, demonstrate, illustrate, summarize, conclude, list, define, report, model, implement, design, consider, involve, simplify, generalize, perform, be based on sg., take into account sg., depend on sg, increase, decrease, evaluate, predict, assign, require, satisfy, ...

**Task:** What is the difference between the following concepts? Give examples when they are used!

evaluate – assess

compute – calculate

derive – infer

approximate – estimate

discover – find

## 2.6 Often needed irregular verbs

The following list contains irregular verbs which are sometimes needed in computer science expressions, excluding the most common ones (which all of you know!):

choose – chose – chosen

find – found – found

hide – hid – hidden

hold – held – held

lead – led – led

lose – lost – lost

rise – rose – risen

seek – sought – sought

show – showed – shown

spin – spun – spun  
split – split – split  
spread – spread – spread  
stick – stuck – stuck

In addition, the last consonant can be doubled before -ed, if

- if the spell is short and stressed: planned, dropped,
- the consonant is 'l': travelled, modelled, biassed

Notice: American English is not so strict, and `ispell` can complain about correct spelling!

## Exercise

Read the given text part and underline useful expressions. Search especially the following kind of expressions:

- Useful verbs and their prepositions in computer science texts.
- How to list advantages or disadvantages without repetition (usually in the beginning of sentences).
- How to compare approaches?
- Any other useful expressions!

The same text is given to two people. Thus, you can discuss with your pair, if you don't understand something. However, it is not important if you don't understand all words.

### 3 Extra material: Tricks for gender-neutral language

<b>Trick</b>	<b>Incorrect</b>	<b>Correct</b>
Use plural	The student returned his solution.	The students returned their solutions
Article "the"		The student returned the solution.
Drop the pronoun	The user himself defines the preferences.	The user defines the preferences.
Special expressions	man, mankind man-machine interface  Researchers' wives mothering chairman Mrs. Smith housewife	people, human beings, humankind user-system interface, human-computer interface Researchers' spouses pareting, nurturing chairperson, chair, head Jane Smith homemaker