How To Write (not only) a Hausarbeit

Ivana Kruijff-Korbayová *

http://www.coli.uni-saarland.de/~korbay
korbay@coli.uni-sb.de

November 10, 2006

Outline of the Talk

- Are you ready to write?

- Writing phases

- Structured writing

- Visual aspects of writing

- Citations and references

- Standard tools

- Final remarks

Ready to Write?

Writing Preparation

- Know the purpose of the paper

  - What is the paper for?
  - Why am I writing it?

- Know what the subject (topic) is

- Know the audience/readers:

  - Who are they?
  - What do they know already?
  - What they need to know?
  - Why will they read it?
  - How will they use the information?

  Note: For "Hausarbeit", think of the audience as yourself before you started working on the subject of the paper.

- Literature research

Literature Research

- The teacher typically assigns you some basic literature

- Don't be satisfied with that!
  - Fill in background, so you understand better.
  - Search for additional relevant work.

Finding literature related to the topic of your paper:

- Top-down: find an overview article or a state-of-art paper to get an overall picture of the issues. Then start deepening out particular aspects.

- Follow up on references in recent paper(s).
  - Library: http://www.coli.uni-sb.de/bib/
  - ACL Anthology: http://acl.ldc.upenn.edu/
  - Citeseer: http://citeseer.ist.psu.edu/
  - Google Scholar: http://scholar.google.com/
  - Google: http://www.google.de/

- Use bibliographies (when available)

Writing Phases

Ready to write . . .

1. Planning phase

2. Drafting phase

3. Revision phase

4. Completion phase

5. Editing phase

6. Finalization phase

DONE !!!

Writing Phases

1. **Plan** (cf. talk by Geert-Jan Kruijff)
   - Determine a concrete *goal*.
   - Decide what kind of paper it is.
   - Decide about the structure of argumentation.
   - Make on *outline* of the paper body.

2. **Draft**: Write a rough draft of the body sections.

3. **Revise**: Make rearrangements, delete or add stuff.

4. **Complete**:
   - Write the conclusions section.
   - Write the introduction section.
   - Write the abstract (executive summary)

5. **Edit**: Read, reformulate, proof-read (spell-check, check consistency).

6. **Finalize**: Add remaining formally required parts.

(Generic) Global Structure

1. **Front matter**: title, author(s), affiliation(s), contact detail(s), table of contents, acknowledgements

2. **Abstract** [5-10 lines]

3. **Introduction section** [cca 1 page]
   (a) Goal statement
   (b) Motivation
   (c) Summary
   (d) Brief outline of structure/sections [1 par.]

4. **Body sections**

5. **Conclusions section** [cca 1/2 page]
   (a) Summary
   (b) Lessons
   (c) Further research

6. **Back matter**
   (a) References
   (b) Appendix

General Writing Tips

- Think the structure through well.

- Use informative (clear and specific) headings.

- Substructure as much as you can (all the way down to sentence level).

- Start each paragraph by a leading topic sentence stating its main point; then develop and substantiate the topic in the remaining sentences.

- Provide road-maps throughout the paper ("guidelines" for the reader): typically, a brief summary at the end of a "theme"/section, or an outline at the beginning of a new one.

Apply the Gricean Maxims

cf. (Grice, 1975)

- Be honest and charitable.

- Include sufficient amount of detail for the readers to understand what is really going on (but do not drown them in details irrelevant to your point).

- Use short sentences; present a single thought, or two related thoughts in each sentence (not more!).

- Make relations between thoughts explicit, for example by using discourse connectives.

- Be concrete and to the point, avoid abstract concepts and formal jargon.
  Example of a Foggy Text:
  Having shown the applicability of analysis of covariance in straightforward research situations, I shall go on to indicate how several other important methodological topics can be profitably conceptualized as isomorphic in logical structure to the general linear model.

  Richard A. Lanham, "Revising Prose"
  (cited by Michael A. Covington)

Modesty in Scientific Writing

- Renounce false modesty, and avoid verbose and convoluted phrasing, such as "It was shown that . . . " or "In this author's opinion . . . "

- Don't be afraid to name the agent or action in a sentence, even when it is "I" or "we"

- *Pluralis modestiae*: besides other uses, first person plural can be used by a single author to refer to him/herself (cf. http://www.rhetorik.ch/Aktuell/ Aktuell_Nov_25_2004.html)

[. . .] scientific research is a group activity and not one's individual merit. I will conform to this custom and mostly use the plural form. I will use the singular form for personal comments and also for particularly strong claims, where I want it to be clear that none of my tutors and colleagues share any part of the blame, might I be proven wrong.   (Sikkel, 1993)[p. 3]

Visual Aspects of Writing

Visual Writing ABC

- Do it, but don't over do it!

- Be consistent in visual formats.

- Use graphics, but always reinforce it with supporting text.

- Use *italics* or **bold-face** to emphasize text.

- Very heavy and very light type are difficult to read.

- Serif font looks and reads better than sans-serif.

- Use 11-12 point font-size for body text.

- Do not overfill pages.

Design Considerations and Conventions

- Think which ideas you could present graphically in a useful way

- Learn from examples, that is, look how other people do it.

- Read typesetting manuals.
  - But different conventions in different languages!
  - For English: The Chicago Manual of Style (CMS, 2003) or Turabian's Manual (Turabian, 1996)

Basic types of visualization:

- Tables
- Charts
- Lists
- Figures
- Examples
- Formulae
- Code

Tables

- Usually, tables are separate from text, and numbered for cross-reference.

- Every table must have a caption and a legend (for abbreviations), and should have explanatory text, so that it can be read on its own.

- A table may be "box-less".

- Use lowercase, at least 10 point type.

- Align numerical data to the decimal point

- Use *N/D* for "no data", *N/A* for "not applicable".

- Leave a space every 5 lines down a column, for readability.

- Be consistent.

- Avoid clutter. Don't repeat the same when you can put it together.

cf. Formatting tables neatly:
http://www.scribe.com.au/tip-w013.html

Charts

- Essential types of charts:
  - Data flow chart
  - Curve chart
  - Column and bar chart
  - Surface (pie) chart

- Use the right chart type for your data:
  - For anything with input→output (e.g., algorithms, system architecture, modules in data processing etc.): data flow chart
  - For parts of the whole: pie or bar and column charts
  - For comparison among items: bar or column chart
  - For trends: curve charts (many periods), column charts (few periods)

Lists

- List types:
  - Numbered (when items follow a certain order)
  - Bulleted (when the items are of equal importance)

- Use sublists when there are too many items (e.g., more than 10 for a numbered list, more then 5 for a bulleted list).

- Sublists can begin with dashes (–), letter numbering or Arabic numbering.

- Each list item

  a. Begins with a capital letter.
  b. Ends with a period if it is a whole sentence, with nothing otherwise.

- Lead-in sentence ends with a colon (:) when it is a sentence, or with nothing, dash (–) or ellipsis (. . .) when it is not a sentence.

Figures

- Use figures for larger examples, bits of code, trees, graphs, system architectures etc.

- Figures are separate from text, and numbered for cross-reference.

- Every figure must have a caption and a legend (for abbreviations), and should have explanatory text, such that it can be understood on its own.

- A figure should appear close to where it is discussed in detail, either before or afterwards. Large figures or figures which are referred to at various places throughout the text are better put into an appendix.

- Always refer to a figure in text, and comment on it / explain it.

Linguistic Examples

Use examples to illustrate the phenomena you are discussing. You can include them

Within text: When I say "You must make the example fit into the sentence structure" I mean that the example should be part of your sentence. You can also typeset it in *italics*.

Between text: Usually you'll need to number them for cross-reference, as illustrated in (1).

(1) Ich   habe   eine   Katze.
     I-Pro/nom have-V/sg Indef/Acc cat-N/Fem
     'I have a cat.'

(http://www.essex.ac.uk/linguistics/clmt/latex4ling/examples/gb4e/)

Always keep examples close to where they are first mentioned. If discussion continues, put it below the example, starting with no indentation. If you return to an example later, use cross-reference. If on a different page, repeat the example.

Separately from text: in a figure (e.g., when example contains very complex annotation)

Other Examples and Code

Give examples whenever they are helpful. Examples make your writing clearer and more precise. You can include them

- Between text: Usually, you'll need to number them for cross-reference.

- Separately from text: in a figure (esp. when they are larger chunks of stuff)
  Note: Put long examples and extensive code (if it needs to be included) into an appendix.

Always comment on / explain examples in text!

Formulae

- Use symbols only when they are helpful.

- Always define all symbols you use, no matter how familiar they may be.

- Explain in text what your formulae mean.

- After a long intervening discussion, it may be needed to redefine less common symbols.

- Avoid using the same symbol for different notions.

- Keep symbol use within the grammatical structure of the sentence.

- Put longer formulae or definitions between text or into figures. Number them for cross-reference.

Abbreviations

- Use abbreviation when needed: e.g., for often repeated terms, for names of theories or approaches/methods.

- Always introduce your abbreviations, e.g.: "Embodied Conversation Agents (ECA) are software programs that . . . " or "the theory of discourse structure introduced by Grosz and Sidner (1986) (henceforth DST)", even if they are common, e.g., DRT.

- Reintroduce abbreviations if needed, e.g., after a long passage of not being used.

- When in doubt, spell it out.

Citations and References

Citations

- Always distinguish between your thoughts and those of others

- Always cite the sources you used for your work!

- Always cite sources of ideas you build on!

- Formulate the crux of the idea using your *own words* and provide a citation. Avoid direct quotations, unless the particular wording is crucial or particularly remarkable.

- If you cite a source about which you know from some other source, but you did not have direct access to it, include a "cited by" citation in the reference (Covington's short sample paper).

- If you cite an unpublished idea person X gave you in discussion, cite it as *(X, p.c.)* –for "personal communication".

Common Citation Formats

a. By **source number** from reference list, e.g.,
   "Convington's sample short paper gives some useful examples of citations and references [1]."

   In this case, keep the reference outside of the grammatical structure of the sentence to make the sentence intelligible without the reference

b. By **author(s) name(s) and publication year**:
   - Outside the grammatical structure of the sentence, e.g.,
     "The sample short paper gives some useful examples of citations and references (Covington 1990)."
   - Within the grammatical structure of the sentence, e.g.,
     "Covington (1990) gives some useful examples of citations and references."
   - Multiple authors: *(X and Y, 1999)* for two authors, *(X et al., 1999)* for more than two authors.

Note: Do not mix formats (a) and (b)!

References

- Every source mentioned in the paper must be listed in the list of references.

- Do not list sources not mentioned in the paper.

- The list of references must be ordered
  a. by source number
  b. by author(s) surname(s)

- Provide all necessary information for each reference.

- There exist different typesetting standards for references defined by various publishers.
  cf. The Chicago Manual of Style (CMS, 2003) `http://www.chicagomanualofstyle.org/` or Turabian's Manual (Turabian, 1996)

- When using LaTeX and BibTeX, you can use a standard style, e.g., *plain* for numbered references; for the named style, e.g., *chicago* or style packages from ACL

References

The information that must be included for each reference depends on the type of the reference. It contains at least

1. Author(s) name(s)
2. Title of the source
3. Publisher
4. Year of publication

Journal or collection articles and papers from conference proceedings contain in addition at least

1. Editor(s) name(s)
2. Title of the journal/volume
3. Issue/Volume identification number
4. Page numbers

Standard Tools

(Some) Standard Tools

- Text editors:
  - Emacs (XEmacs) and other ASCII editors (any platform)
  - MS Word and other editors (under Windows)

- Text processing (typesetting) tools/programs:
  - LaTeX and BibTeX (any platform)

- Graphics editors for drawing simple pictures:
  - xfig (Unix, Linux)
  - MS Draw (Windows)

- Spelling checkers (never forget to use one!)
  - Esc-x ispell in Emacs
  - F7 (?) in MS Word

- Versioning control: keeping track of changes
  - CVS (Unix, Linux)
  - "trace changes" in MS Word

- . . .

LaTeX

- Professional typesetting "programming language" (Goossens et al., 1994)

- Especially designed for typesetting "difficult texts"

- Assumed standard in computational linguistics (and many other fields, though not in humanities)

- Standard *article, report, book, letter* etc. styles

- Many other more or less standard styles provided by institutes, publishers, etc.

- Easy to change style(s)

- Standard packages for easier typesetting of, e.g., linguistic examples, graphics, trees, tables, etc.

- Standard commands for including and styles for formatting citations and references

- Additional tool for creating a reference list from a database of bibliography items (BibTeX) —this way, one does not always re-create the references

---

MS Word

- "What you see is what you (hope to) get"

- Easier for beginners (just start typing)

- Typesetting texts with formal stuff gets difficult

- Assumed standard in the humanities, but (I think) also in most of the commercial world

- Standard templates for basic document types

- Publishers usually also provide templates

- Changing a template or already used styles does not alway work well

- Putting together contributions from various sources usually not without problems

- MS tools for graphics

- Support handling citations and references: EndNotes (I think)

Wrapping Up

Avoid Common Mistakes

Make sure your paper contains:

- all important structural parts
- a clear goal definition
- motivations for why something is done and why it is done that way
- road-maps (reader "guidelines") throughout
- examples and/or concrete illustrations
- tables and graphics when suitable
- a discussion expressing own critical assessment
- clear conclusions

Make sure your paper is formally correct:

- complete and correct citations and references
- correct cross-references
- properly introduced abbreviations
- consistent terminology and formatting
- correctly spelled words, grammatical sentences
- etc.

See handout for serious tips put in a humorous way!

Further Reading Suggestions

- (Beer, 1992) contains also other useful material.

- Presentation about technical writing in German:
  `http://www.math-inf.uni-greifswald.de/informatik/koehler/`
  `publiz/studenten/techschr.ppt`

- Check books on writing and speaking in the library.

- Check materials on the website of the Society for Technical Communication `http://www.stc.org/` e.g., conference session materials, seminars, etc.

- Type "technical writing guide" in *google*

- Check the Chicago Manual of Style (CMS, 2003) or Turabian's Manual (Turabian, 1996)

- For tools available at CoLi and basic tips on using them, check first of all the *sysadmin* pages

Just Do It. Good Luck!

References

David F. Beer, editor. 1992. *Writing and Speaking in Technology Professions A Practical Guide*. IEEE.

H. P. Grice. 1975. Logic and conversation. In P. Cole and J. Morgan, editors, *Syntax and semantics*, number 3. Academic Press, New York.

Kate Turabian. 1996. *A Manual for Writers of Term Papers, Theses, and Dissertations, 6th edition*. Chicago: University of Chicago Press.

University of Chicago Press Staff, editor. 2003. *The Chicago Manual of Style, 15th Edition*. Chicago: University of Chicago Press.

Michael Goossens, Frank Mittelbach, and Alexander Samarin. 1994. *The LaTeX Companion*. Addison Wesley.

Nicolaas Sikkel. 1994. *Parsing schemata*. Ph.D. disertation, Twente University, Enschede, the Netherlands.