Bachelorarbeit

im Studiengang Computerlinguistik

an der Ludwig-Maximilians-Universität München

Fakultät für Sprach- und Literaturwissenschaften

Your Thesis Title

vorgelegt von
Your Name

Betreuer:	Your Supervisor's Name
Prüfer:	Your Examiner's Name
Bearbeitungszeitraum:	Start of Registration – Submission Deadline

Abstract

This abstract should be written in the language of the thesis and must not exceed 250 words. It should provide a concise summary of the work, clearly stating the research topic, objectives, methodology, and the main results or contributions.

Contents

1. Using This Template

1.1. Getting Started

This document serves as a template for Bachelor's and Master's theses in Computational Linguistics at the Center for Information and Language Processing (CIS, LMU Munich). It is designed to be used in either English or German and reflects the formatting, structure, and layout expected for submission. It includes pre-set formatting, title page, citation setup, bibliography, and useful examples for figures, tables, equations, and more.

All content is meant to be replaced by your own work. The sections and examples are included to show LaTeX usage and best practices. You may remove this entire chapter from your final submission.

1.2. Thesis Layout

A typical thesis consists of several numbered chapters, each structured into sections and subsections. You can also use unnumbered sections or paragraphs when appropriate. Use `\chapter{}`, `\section{}`, `\subsection{}`, and `\paragraph{}` to organize your content. Below are some examples:

1.2.1. Example Subsection

This is a numbered subsection.

Unnumbered Subsection

This subsection does not appear in the table of contents.

Titled Paragraph This is a titled paragraph.

1.2.2. Example Chapter Outline

While the structure of your thesis depends on your topic and supervisor preferences, a common organization in Computational Linguistics is:

1. **Introduction** – Introduce your topic, motivation, and research question(s).

2. **Background** – Present related work and theoretical or methodological foundations.

3. **Methodology** – Describe your approach, data, and methods used.

4. **Analysis** – Show your experiments, data analysis, and results.

5. **Discussion** – Summarize findings in a conclusion, discuss limitations, and outline future work.

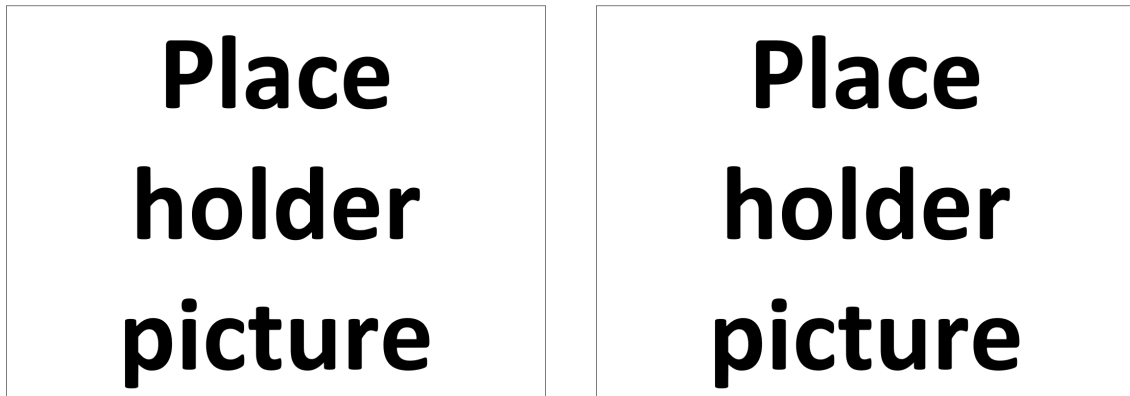This structure is flexible – adapt it to your thesis content and goals.

Figure 1.1.: Example of side-by-side images.

1.3. Citations

This template uses the `natbib` package[1] with the `acl_natbib` bibliography style[2].

You can use the command `\citet` (textual citation) to get "author (year)" style, like this citation to a paper by Brown et al. (2020). Use `\citep` (parenthetical citation) for "(author, year)" citations (Brown et al., 2020). Use `\citealp` (no parentheses) to get "author, year", which is useful inside your own parentheses, e.g., (see Brown et al., 2020).

- `\citet{brown2020gpt3}` → Brown et al. (2020)   (textual citation)
- `\citep{brown2020gpt3}` → (Brown et al., 2020)   (parenthetical citation)
- `\citealp{brown2020gpt3}` → Brown et al., 2020 (author-year without parentheses)
- `\citeyearpar{brown2020gpt3}` → (2020)   (year only, in parentheses)
- `\citeauthor{brown2020gpt3}` → Brown et al.   (author name only)

To cite multiple sources at once, separate the keys with commas inside `\citep`:

- `\citep{brown2020gpt3,vaswani2017attention}` → (Brown et al., 2020; Vaswani et al., 2017)

Note that in this setup, the basic command `\cite{...}` behaves like `\citet{...}` – it produces a textual citation.

1.4. Figures, Tables, and Equations

1.4.1. Figures

See Figure 1.1 for an example of a figure and its caption. If you include third-party images, clearly credit the source in the caption or a footnote.

Images used in this template are located in the `img/` folder. You should place your own images there as well, or organize them into subfolders if needed for clarity.

1.4.2. Tables

This template uses the `booktabs` package for cleaner and more professional table formatting. It replaces heavy horizontal lines (`\hline`) with typographically correct rules like `\toprule`, `\midrule`, and `\bottomrule`.

---

[1] https://ctan.org/pkg/natbib
[2] https://github.com/acl-org/acl-style-files

See Table 1.2 for an example of a standard table and its caption. You can create tables manually or use an online tool like https://www.tablesgenerator.com/ for convenience.

For tables that span the full text width and contain longer or wrapped text entries, use the `tabularx` environment in combination with `booktabs`. This is especially useful when one or more columns contain flexible-width content. See Table 1.1 for an example.

Field	Description
`title`	The full title of the referenced work
`author`	The author(s) of the work
`year`	The year of publication

Table 1.1.: Example of a table using `tabularx` for flexible layout.

1.4.3. Placement of Figures and Tables

LaTeX will automatically place figures and tables where they fit best, typically at the top or bottom of a page. This helps maintain consistent and readable layout.

To prevent floating and force placement exactly where defined, use `[H]` – enabled by the `float` package included in this template. See Table 1.1 for an example.

1.4.4. Writing Captions

Captions for figures and tables should be informative and provide enough detail to be understood on their own, without needing to read the full text. A good caption describes what the figure or table shows and how it relates to the content in the text.

Captions may consist of full sentences or short explanatory phrases, especially when the figure or table presents complex information. For figures with multiple elements (e.g., subfigures, multi-column tables), captions can span multiple sentences.

1.4.5. Equations

Equations should be centered and numbered, see Equation 1.1 for an example.

$$A = \pi r^2 \tag{1.1}$$

For short expressions within the text, use inline math mode by wrapping the content in `$...$`, e.g. πr^2. You may also see `$$...$$` used for display math without numbering. However, this syntax is not recommended in LaTeX – instead, use environments like `equation` or `align` for display math.

1.4.6. Referencing

All elements – chapter, (sub)sections, figures, tables, equations – can be referenced using `\label` and `\ref`, as in: "This is Section 1.4.6 and Figure 1.1."

1.5. Special Characters and Encoding

This template supports direct input of accented characters (e.g., ä, ö, é) using UTF-8 and T1 font encoding. However, other characters should be written using LaTeX commands to avoid encoding issues, such as dashes (use `--` or `---`), quotation marks (use `` `` `` and `''` or `\enquote`), and math symbols (e.g., `\in`, `\sum`). Avoid copy-pasting symbols from websites or PDFs, as they may introduce hidden encoding problems.

Command	Output	Command	Output
{\"a}	ä	{\c c}	ç
{\^e}	ê	{\u g}	ğ
{\`i}	ì	{\l}	ł
{\.I}	İ	{\~n}	ñ
{\o}	ø	{\H o}	ő
{\'u}	ú	{\v r}	ř
{\aa}	å	{\ss}	ß

Table 1.2.: Example commands for accented characters, to be used in Bibtex entries.

1.6. Bibliography

All references should be listed in the `references.bib` file using Bibtex format. Cite sources in the text using commands like `\citep` and `\citet` (see Section 1.3).

The bibliography is automatically generated at the end of the document. By default, the template includes all entries in `references.bib` using the `\nocite{*}` command. If you want to include only the works you actually cite, comment out or remove that line.

Avoid using Unicode characters in Bibtex entries – they can break compilation or alphabetization. Use LaTeX-style accent commands instead, as shown in Table 1.2.

When available, include DOIs or URLs in your entries. Use the `doi` field for DOIs and the `url` field for links. If these are present, the paper title in the bibliography will be hyperlinked (via the `hyperref` package).

1.7. Appendices

Supplementary Material

These appendices are optional and can be used to include one or more chapters of supplementary material that supports your thesis but is not part of the main chapters. Typical examples include:

- Annotation guidelines

- Extended tables or figures

- Evaluation details

- Additional data or code documentation

Appendix chapters are labeled with letters (A, B, C...) and should follow the main content of your thesis. An example is already included in the template.

Submitted Software and Data Files

A separate mandatory appendix (also included in the template) is provided for describing the contents of any supplementary software or data you are submitting along with your thesis, such as source code, datasets, and documentation.

References

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems 33*, pages 1877–1901.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1*, pages 4171–4186.

Bonan Min, Hayley Ross, Elior Sulem, Amir Pouran Ben Veyseh, Thien Huu Nguyen, Oscar Sainz, Eneko Agirre, Ilana Heintz, and Dan Roth. 2023. Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Computing Surveys*, 56(2):1–40.

OpenAI. 2022. Introducing ChatGPT. Accessed: 2025-03-01.

OpenAI. 2024. GPT-4 technical report. *arXiv*, abs/2303.08774.

Yubing Ren, Yanan Cao, Ping Guo, Fang Fang, Wei Ma, and Zheng Lin. 2023. Retrieve-and-sample: Document-level event argument extraction via hybrid retrieval augmentation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Volume 1*, pages 293–306.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30*.

List of Figures

List of Tables

A. Supplementary Material

B. Submitted Software and Data Files

Example:

- `code/` - Python scripts used for model training and evaluation

- `data/` - Preprocessed dataset and annotation guidelines

- `thesis.pdf` - PDF version of this thesis

- `README.md` - Instructions for setting up the environment

Selbstständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig angefertigt, alle Zitate als solche kenntlich gemacht sowie alle benutzten Quellen und Hilfsmittel angegeben habe.

München, den TT.MM.JJJJ

. . . .
Your Name