

Research Master Thesis

# New Insights in Computational Lexicology

Johnny B. Good

*a thesis submitted in partial fulfilment of the  
requirements for the degree of*

**MA Linguistics**  
(Human Language Technology)

**Vrije Universiteit Amsterdam**

Computational Lexicology and Terminology Lab  
Department of Language and Communication  
Faculty of Humanities



Supervised by: Joan Smith, Jan de Vries  
2<sup>nd</sup> reader: Summer Reads

Submitted: June 30, 2020



# Abstract



# Declaration of Authorship

I, John Bumblebee Good, declare that this thesis, titled *New Insights in Computational Lexicology* and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Date:

Signed:



# Acknowledgments





# List of Figures



# Contents

<b>Abstract</b>	<b>i</b>
<b>Declaration of Authorship</b>	<b>iii</b>
<b>Acknowledgments</b>	<b>v</b>
<b>List of Figures</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Latex tips</b>	<b>3</b>
2.1 Latex resources . . . . .	3
2.2 Structure of a $\text{\LaTeX}$ project . . . . .	3
2.3 Citations . . . . .	4
2.3.1 The natbib package . . . . .	4
2.3.2 Citing conventions . . . . .	5
2.3.3 Bibliography management . . . . .	5



# Chapter 1

## Introduction

This document provides a template for Master theses at the CLTL, as well as a number of L<sup>A</sup>T<sub>E</sub>X tips, which are presented in chapter 2.

To use this template as a starting point for writing your thesis:

- fill in the appropriate fields (name, title, etc.) in `mathesis.tex`;
- fill in the file stubs in the `tex` folder: `abstract.tex`, `acknowledgments.tex`, etc.
- replace `bib/example.bib` by your own `bib` file(s);
- and anything else needed to make this template your thesis. All the best!



## Chapter 2

# Latex tips

This chapter provides tips for using L<sup>A</sup>T<sub>E</sub>X for writing your thesis, as well as more general tips for bibliographical references.

### 2.1 Latex resources

Latex is extremely well documented. The following resources will all give you an easy-to-step-in introduction, and an extensive reference to L<sup>A</sup>T<sub>E</sub>X:

- the *Not so short introduction to L<sup>A</sup>T<sub>E</sub>X2 $\epsilon$* : <https://tobi.oetiker.ch/lshort/lshort.pdf>
- the L<sup>A</sup>T<sub>E</sub>X wiki book: <https://en.wikibooks.org/wiki/LaTeX>
- the Overleaf documentation: <https://www.overleaf.com/learn>

Additionally, the *Comprehensive L<sup>A</sup>T<sub>E</sub>X symbols list* is worth a booktab. Finally, L<sup>A</sup>T<sub>E</sub>X has a lot of packages to offer for additional functionality, all stored on CTAN: <https://www.ctan.org>.

### 2.2 Structure of a L<sup>A</sup>T<sub>E</sub>X project

Your thesis may run into more than 50 pages, and include as many pictures. It is therefore recommended to structure your L<sup>A</sup>T<sub>E</sub>X thesis file into parts (a natural division consists in keeping a separate file for each chapter).

For this template, the file `mathesis.tex` is the main file. It links to content files stored in the `tex` folder. You will find there files for, e.g., the abstract and acknowledgments, but you can also add your chapter files. Likewise, you can store images in the `img` folder.

L<sup>A</sup>T<sub>E</sub>X documents can be included into one another using the `\include` command: in the main file `mathesis.tex`, the assertion `\include{tex/abstract}` looks for the file `tex/abstract.tex` and inserts its content into `mathesis.tex`.

## 2.3 Citations

### 2.3.1 The natbib package

The `natbib` package allows to refer to BibTeX bibliographical references and format them for insertion in a L<sup>A</sup>T<sub>E</sub>X document. BibTeX bibliography items are stored in a `.bib` file.

For instance, the example bibliography `./bib/example.bib` contains two entries:

```
@inproceedings{sommerauer-etal-2019-towards,
  Address = {Wroclaw, Poland},
  Author = {Sommerauer, Pia and Fokkens, Antske and Vossen, Piek},
  Booktitle = {Proceedings of the 10th Global Wordnet Conference},
  Pages = {85--95},
  Title = {Towards Interpretable, Data-derived Distributional
Semantic Representations for Reasoning: A Dataset
of Properties and Concepts},
  Url = {https://clarin-pl.eu/dspace/handle/11321/718},
  Year = {2019},
  Bdsk-Url-1 = {https://clarin-pl.eu/dspace/handle/11321/718}}

@inproceedings{van-aggelen-etal-2019-larger,
  Address = {Turku, Finland},
  Author = {van Aggelen, Astrid and Fokkens, Antske and Hollink,
Laura and van Ossenbruggen, Jacco},
  Booktitle = {Proceedings of the 22nd Nordic Conference on
Computational Linguistics},
  Pages = {44--54},
  Publisher = {Link{"o"}ping University Electronic Press},
  Title = {A larger-scale evaluation resource of terms and
their shift direction for diachronic lexical semantics},
  Url = {https://www.aclweb.org/anthology/W19-6105.pdf},
  Year = {2019},
  Bdsk-Url-1 = {https://www.aclweb.org/anthology/W19-6105.pdf}}
```

The first line of each entry provides a label for references: *sommerauer-etal-2019-towards*, *van-aggelen-etal-2019-larger*. These labels can be referred to in the L<sup>A</sup>T<sub>E</sub>X document to provide formatted bibliographical references.

The two most commonly employed commands are `\cite` (or equivalently `\citet`) and `\citep`. For instance, `\citet{sommerauer-etal-2019-towards}` will appear as Sommerauer et al. (2019), while `\citep{sommerauer-etal-2019-towards}` will appear as (Sommerauer et al., 2019).

You can cite several papers with a single citation. For instance, the command `\citep{sommerauer-etal-2019-towards,van-aggelen-etal-2019-larger}` results in (Sommerauer et al., 2019; van Aggelen et al., 2019).

See the Natbib package documentation or the usual L<sup>A</sup>T<sub>E</sub>X references for more information.



### 2.3.2 Citing conventions

It is convention to integrate the name of the authors in the text as much as possible, and to use `\cite` as only the year of the reference is then parenthesized. The `\citep` can be used when the name of the authors is not directly part of the sentence.

For instance, you would use `\cite` for “the work of van Aggelen et al. (2019)”, and `\citep` for “...research on diachronic lexical semantics (van Aggelen et al., 2019)”.

### 2.3.3 Bibliography management

We recommend that you use a bibliography management tool to edit `bib` files, like BibDesk for Mac, or JabRef.

This will provide you with a better overview of your bibliography as it grows, while facilitating the addition of new entries—this is as simple as copying the `bibtex` reference of an article and pasting it in the *bibtex source* field of your management tool.

Note that you can refer to distinct `bib` files in a  $\text{\LaTeX}$  document. Suppose for instance that you would like to keep apart references from the background chapter and from the other chapters, in `background.bib` and `research.bib`. You can collect both files with: `\bibliography{bib/background,bib/research}`.

Note however that  $\text{\LaTeX}$  will issue a warning if citations overlap between `bib` files.



# Bibliography

- P. Sommerauer, A. Fokkens, and P. Vossen. Towards interpretable, data-derived distributional semantic representations for reasoning: A dataset of properties and concepts. In *Proceedings of the 10th Global Wordnet Conference*, pages 85–95, Wroclaw, Poland, 2019. URL <https://clarin-pl.eu/dspace/handle/11321/718>.
- A. van Aggelen, A. Fokkens, L. Hollink, and J. van Ossenbruggen. A larger-scale evaluation resource of terms and their shift direction for diachronic lexical semantics. In *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, pages 44–54, Turku, Finland, 2019. Linköping University Electronic Press. URL <https://www.aclweb.org/anthology/W19-6105.pdf>.