Universität zu Köln

Faculty of Arts and Humanities, Department of Linguistics

Master's Thesis

TITLE OF THE THESIS

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

Master of Arts

in the

Master Linguistik

Author: AUTHOR NAME

Matriculation Number: MATRICULATION NUMBER

Supervisor: SUPERVISOR NAME

DATE OF SUBMISSION

Table of Contents

List of Figures

List of Tables

1 Introduction

Write your introduction text here. Name your sections (e.g. "#sec-introduction"). You can use the section names for cross references like this:

"In Section 4, …"

2 Chapter Title

Add some text about chapter 2, if you like.

2.1 Subchapter Title

Add some text about chapter 2.1, if you like.

2.1.1 Subchapter Title

You can include references (via BibTex from the file `references.bib`) and images (from the `image` folder) in your document.

Citations are inserted using their BibTeX keys in square brackets. Using `@key` includes the author's name in the citation, while `-@key` suppresses the author name and is useful when the author is already mentioned in the text. Page numbers can be added after a colon.

Figures are numbered automatically and can be referenced in the text using their figure labels (e.g. `@fig-dialects`).

For example:

"Cornelissen (2015: 73) refers to North Rhine-Westphalia as *Land der Tausend Dialekte* 'land of a thousand dialects' to emphasise the federal state's rich linguistic diversity. This diversity is illustrated in Cornelissen's map shown in Figure 1, published by the LVR-Institut für Landeskunde und Regionalgeschichte (2020, July 22), which categorises the state into ten distinct dialect areas. The map's structure is based on 19th-century dialect research (Cornelissen, 2015: 73)."

2.1.2 Subchapter Title

If necessary, you can add another pagebreak anywhere in the document like this (try deleting the page break to see how the pdf document comes out without it):

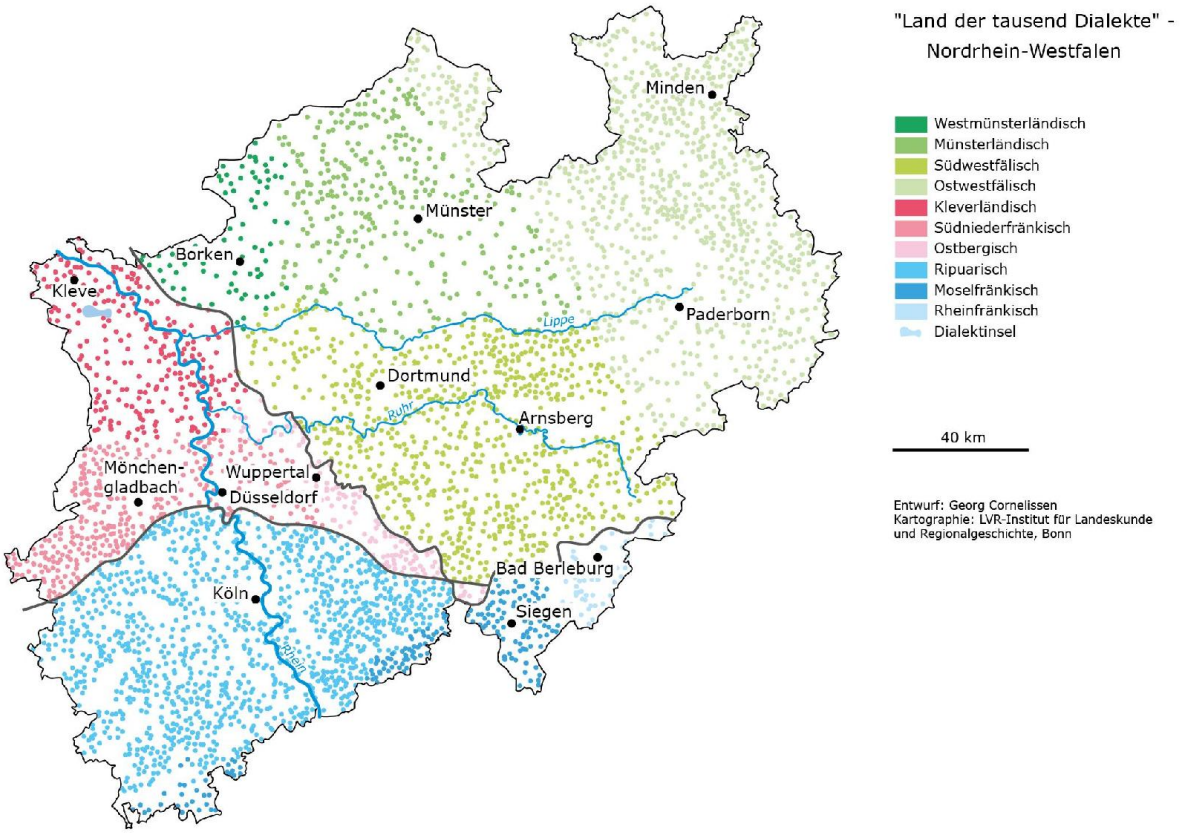

Figure 1: Dialects in NRW (LVR-Institut für Landeskunde und Regionalgeschichte, 2020, July 22)

2.2 Subchapter Title

If your code creates a plot that you want to include and refer to as a figure (e.g. visualizations created with `ggplot` in your results section), you can do it like this:

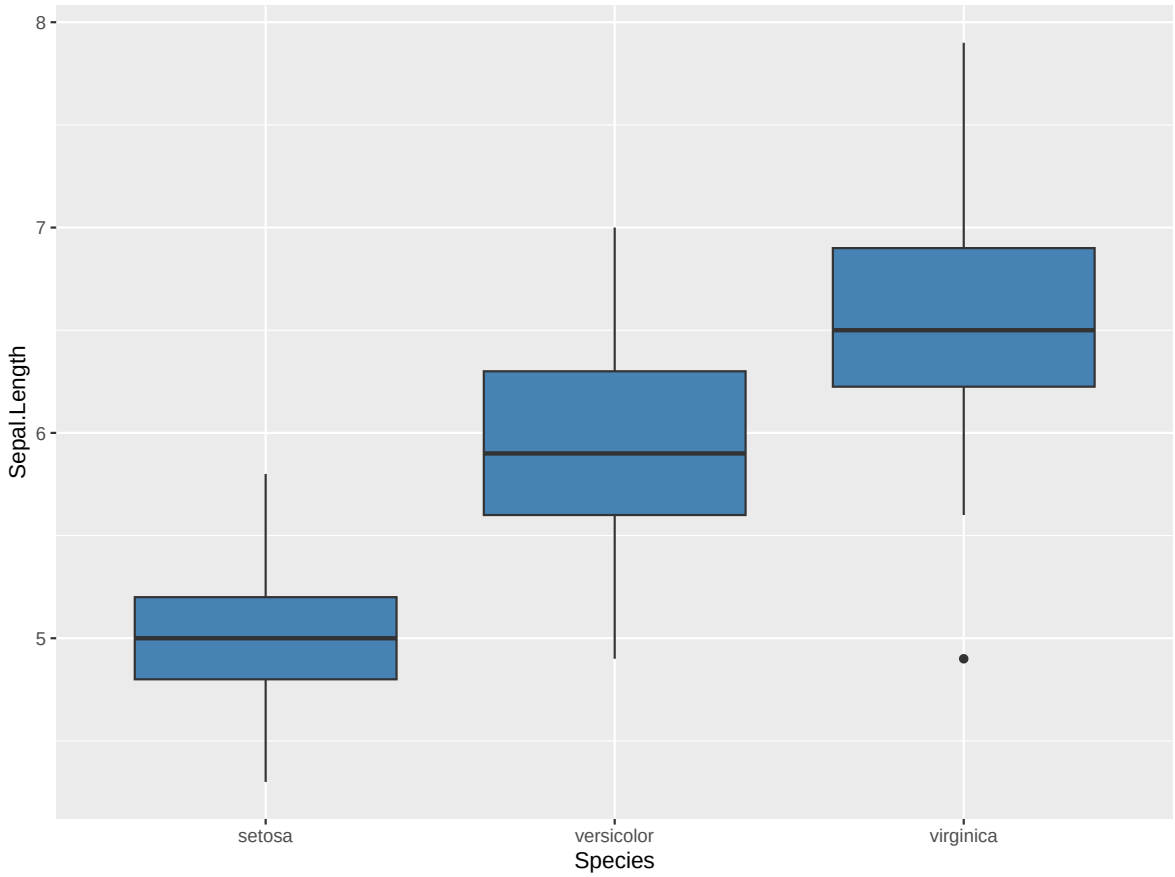

Figure 2: Sepal Length by Species

Setting `echo = FALSE` means that the R code itself is not shown in the rendered document and only the resulting plot is displayed. The text below the code chunk is displayed as the figure caption. The figure can be referenced in the text using Figure 2.

2.2.1 Subchapter Title

Tables can be created directly in RStudio by selecting "Table" → "Insert Table…" from the menu bar. Tables defined in this way can be cross-referenced within the text using Quarto's table labels. For example:

"As illustrated in Table 1, …"

Table 1: Morphological Variables with Examples

Linguistic feature	**Variable**	**Description**	**Examples**
diminutive suffix	biss (chen/ken)	*-chen* vs. *-ken* in lexical variants of *Schnapsglas* 'shot glass'	*Pinnchen*; *Pinneken*; *Schnapsgläschen*; *Schnapsgläsken*
	pinn (chen/ken)	*-chen* vs. *-ken* in lexical variants of *bisschen* 'a little'	*bisschen*; *bissken*; *Stückchen*; *Stücksken*

Alternatively, tables can be generated from dataframes. One approach is to use the `kable()` function from the {knitr} package, which converts data frames into pretty tables for pdf output. This is an example of this approach using the built-in `iris` dataset:

"Table 2 shows…"

Table 2: Example Table Showing the First 6 Rows of the Iris Dataset

Sepal.Length	Sepal.Width	Petal.Length
5.1	3.5	1.4
4.9	3.0	1.4
4.7	3.2	1.3
4.6	3.1	1.5
5.0	3.6	1.4
5.4	3.9	1.7

2.2.2 Subchapter Title

To create a footnote[1], head to "Insert" in RStudio and choose "Footnote".

2.3 Subchapter Title

You can even include IPA symbols/transcriptions like this:

"The *Benrather Linie* separates German dialects which have retained the plosive [k] (as in *maken* 'to make') from those where [k] shifted to [ç] (as in *machen* 'to make') during the Second

[1] Add some footnote text.

Sound Shift, a process that unfolded in stages, starting before the 7th century (Cornelissen, 2015: 21-22; Schmidt and Möller, 2019: 526). It also marks other phonological contrasts, such as [p] vs. [f] and [t] vs. [s], though not consistently across all lexical items. North of the *Benrather Linie* lies the Low German dialect continuum, while south of it are the Central German dialects (Cornelissen, 2015: 21-25)."

2.3.1 Subchapter Title

Include language examples like this:

"The use of *sein* 'to be' instead of *haben* 'to have' as the auxiliary in the present perfect is a regional feature in northwestern Germany, including Westphalia (Elspaß and Möller, 2003) (round 4, question 1c). For example, in the following sentences, examples 1.a and 2.a use a form of *haben*, while example 1.b and 2.b use a form of *sein*, with no change in meaning:

(1) (a) *Ich habe mit Krafttraining angefangen.*
'I (have) started strength training.'

(b) *Ich bin mit Krafttraining angefangen.*
'I (have) started strength training.'

(2) (a) *Ich habe angefangen zu schreiben.*
'I (have) started writing.'

(b) *Ich bin angefangen zu schreiben.*
'I (have) started writing.' "

2.3.2 Subchapter Title

You can also use R code directly in the text (called "inline code") to dynamically display values computed in your code. For example, the following code chunk defines a value:

You can then insert that value into the text using inline R code, so that it updates automatically if the code changes. For example:

"I included the sociodemographic metadata for the 42 participants."

Inline R code can also be used for more complex code, a simple example: 43

2.3.3 Subchapter Title

If you want to get a bit nerdy, you can include LaTeX formulas and inline math:

"The distance between two individuals is mathematically expressed as:

$$d^2_{i,i'} = C \sum_{k=1}^{K} \frac{(x_{ik} - x_{i'k})^2}{I_k}$$

This formula calculates the squared distance to determine which categories the individuals share or differ on: for each category k, we examine whether individual i and individual i' possess that category (with values 1 or 0), and compute the squared difference of their values $(x_{ik} - x_{i'k})^2$. The squared difference equals 0 if both either have or do not have the category, and 1 if they differ. Since not all categories are equally informative, these differences are weighted by the frequency of the respective category, i.e. by dividing by I_k (the total number of individuals possessing category k). This way, uncommon but distinctive features contribute more to distinguishing individuals (Husson et al., 2017: 134)."

3 Chapter Title

Add your chapter text.

4 Discussion

Add your discussion text.

5 Conclusion

Add your conclusion text.

6 References

6.1 Literature

Cornelissen, G. (2015). *Kleine Sprachgeschichte von Nordrhein-Westfalen*. Greven Verlag Köln.

Elspaß, S., and Möller, R. (2003). *Atlas zur deutschen Alltagssprache (AdA)*. https://www.atlas-alltagssprache.de. Accessed: 2025-08-02.

Husson, F., Le, S., and Pagès, J. (2017). *Exploratory Multivariate Analysis by Example Using R* (2nd ed.). Chapman; Hall/CRC. https://doi.org/10.1201/b21874

LVR-Institut für Landeskunde und Regionalgeschichte. (2020, July 22). *NRW - "Land der tausend Dialekte"*. https://dat-portal.lvr.de/orte/dialektkarten/einteilungskarten/nrw-land-der-tausend-dialekte.

Schmidt, J. E., and Möller, R. (2019). Historisches Westdeutsch/Rheinisch (Moselfränkisch, Ripuarisch, Südniederfränkisch). *Sprache Und Raum. Ein internationales Handbuch der Sprachvariation*, *4*, 515–550. https://doi.org/10.1515/9783110261295-005

6.2 Packages Used

```
R version 4.5.1 (2025-06-13)
Platform: aarch64-apple-darwin20
Running under: macOS Tahoe 26.2

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRlapack.dylib;

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

time zone: Europe/Berlin
tzcode source: internal

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
[1] knitr_1.50    ggplot2_4.0.0

loaded via a namespace (and not attached):
 [1] vctrs_0.6.5        cli_3.6.5          rlang_1.1.6        xfun_0.53
 [5] generics_0.1.4     S7_0.2.0           jsonlite_2.0.0     labeling_0.4.3
```

```
 [9] glue_1.8.0         htmltools_0.5.8.1  scales_1.4.0       rmarkdown_2.29
[13] grid_4.5.1         evaluate_1.0.5     tibble_3.3.0       fastmap_1.2.0
[17] lifecycle_1.0.4    compiler_4.5.1     dplyr_1.1.4        RColorBrewer_1.1-3
[21] pkgconfig_2.0.3    rstudioapi_0.17.1  farver_2.1.2       digest_0.6.37
[25] R6_2.6.1           tidyselect_1.2.1   pillar_1.11.0      magrittr_2.0.4
[29] withr_3.0.2        tools_4.5.1        gtable_0.3.6
```

6.3 Package References

[1] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2025. https://www.R-project.org/.

[2] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN: 978-3-319-24277-4. https://ggplot2.tidyverse.org.

[3] H. Wickham, W. Chang, L. Henry, et al. *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. R package version 4.0.0. 2025. https://ggplot2.tidyverse.org.

[4] Y. Xie. *Dynamic Documents with R and knitr*. 2nd. ISBN 978-1498716963. Boca Raton, Florida: Chapman and Hall/CRC, 2015. https://yihui.org/knitr/.

[5] Y. Xie. "knitr: A Comprehensive Tool for Reproducible Research in R". In: *Implementing Reproducible Computational Research*. Ed. by V. Stodden, F. Leisch and R. D. Peng. ISBN 978-1466561595. Chapman and Hall/CRC, 2014.

[6] Y. Xie. *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.50. 2025. https://yihui.org/knitr/.

Appendix A: Additional Tables

Add additional tables, if you like.

Appendix B: Additional Figures

Add additional figures, if you like.

Appendix C: Code

The complete code used for the analyses in this thesis is contained within the Quarto (`.qmd`) file, which is accessible via the following Git repository: LINK