How to write a paper in linguistics

Guidelines for writing scientific papers (*Schriftliche Arbeit im Selbststudium*) and degree theses
in Comparative Linguistics and Evolutionary Linguistics
at the Department of Comparative Language Science, University of Zurich

English translation of the original document (February 27th, 2019)

PRELIMINARY VERSION

August 28th, 2020

1 Introduction

This document summarizes the guidelines for writing term papers and degree theses in comparative linguistics at the University of Zurich. In its present form, it contains rules, guidelines, and recommendations, and is based on previous guidelines compiled by F. Zuñiga and M. Jenny (*Wie schreibt man eine Arbeit*, FS 2007), B. Bickel (*Hinweise zum Schreiben von Hausarbeiten*, 2011), and R. Schikowski as well as references to M. Haspelmath's *The Generic Style Rules* (accessed on June 17th, 2020) (http://www.academia.edu/7370927/The_Generic_Style_ Rules_for_Linguistics).

2 Formalities

2.1 Word Processing Software

The choice of word processing software is free. However, the use of LaTeX is recommended, because it provides the following benefits:

- platform independency
- maximum stability of formatting and contents
- referencing database integration
- text-based format can be accessed with terminal-tools and scripting languages
- high standard of typographical layout

In order to be able to enter Unicode directly into the source code, XeLaTeX should be used for compilation. Bibliographic databases in BibTeX format can be created with BibDesk or exported from countless other applications (Zotero, Google Scholar…). For LaTeX beginners, the department offers a template that contains a number of important commands and just needs to be adapted. Moreover, the UZH regularly offers introductory LaTeX courses.

2.2 Cover Page

The first page of the paper is a cover without pagination. For term papers and theses, it contains the following information (see template on the following page):

- University of Zurich, Department of Comparative Language Science
- Title of the paper
- Paper type, major/minor of studies, ECTS total
- Hand-in date (month and year)
- Author
- Instructor or supervisor

Department of Comparative Language Science

Raising and control in Kiranti languages

BA-thesis in Comparative Linguistics

Major (90 ECTS)

Submitted in January 2020

Author: Henry McCoy

Supervisor: Charles Xavier

2.3 Formatting

The individual **word count** has to correspond to the paper topic and is set by the supervisor or instructor. The following benchmark figures serve as reference:

- Term paper (*Schriftliche Arbeit im Selbststudium*), 3 ECTS: 4,000 – 6,000 words
- Term paper (*Schriftliche Arbeit im Selbststudium*), 6 ECTS: 8,000 – 12,000 words
- Bachelor thesis (9 / 15 ECTS): 10,000 – 17,000 words / 14,000 – 26,000 words
- Master thesis (30 ECTS): 29,000 – 52,000 words

The following **formatting** standards have to be applied:

- A4 page set up
- Margins: 2.5 cm or LaTeX default
- Font size: 11 pt
- Single line spacing
- Page numbers (with the exception of the cover)

Fonts should be neutral in keeping with the objective style of the text itself:

- Fonts should remain the same within the text body, and only be changed to carry structural meaning (e.g. for headings, interlinear glossing etc.).
- Avoid playful or hard-to-read fonts. For printing, serif fonts that include as many Unicode types as possible are recommended (in particular IPA and diacritics). Popular fonts are e.g. Times New Roman, Linux Libertine, Noto, Junicode (especially for complex diacritics combinations), Gentium, or DejaVu. The UZH's home font, Palatino, and the LaTeX standard font, Computer Modern, unfortunately do not contain all italicized IPA letters.
- Colored font is not permitted. Bold and italicized font are to be avoided as much as possible, if their use does not follow established linguistic conventions (see Section 3.4).

Pages should be **numbered** throughout. Well-considered **footnotes** are permitted. Endnotes should be avoided, because they split up information that belongs together. Footnotes should be used in the form of superscript figures following punctuation (i.e. *last word.*[1], not *last word*[1].).

The formatting of the paper should be carefully rendered as clean and readable as possible:

- The text body should be structured into paragraphs with line breaks. Avoid long paragraphs for reasons of readability.
- Tables and figures should be numbered and provided with a caption. In referring to these elements, the words *Table* and *Figure* need to be capitalized in the text (e.g. *Table 1*, *Figure 2.3*).
- Table and figure captions should be as explicit and self-explanatory as possible, and have to contain the units of measurement represented in the table/figure. Figure legends can be added to the figure itself or spelled out as part of the caption.
- Avoid page breaks in continuous units (e.g. examples with glossing spread out over multiple pages).

2.4 Paper Submission

Term papers have to be handed in as an electronic PDF file. WYSIWYG-files such as .doc or .odt cannot be accepted due to the instability of their formatting. Large tables or glossed texts have to be added as an appendix, which needs to be submitted alongside the paper (but not printed out). Suitable table file types are .csv or .tab, suitable file types for glossing are .xml (ELAN) or .txt (Toolbox, Shoebox). Program-specific file types (.rda, .xls etc.) should be avoided.

3 Scientific Writing

3.1 Structure

The great majority of publications in empirical sciences follows a default structure as described below. This structure is also recommended for papers in comparative linguistics, and is a requirement for papers with a phonetic, quantitative-typological, corpus linguistic, psycholinguistic, and animal communication/cognition focus. Even if papers in other fields do not adopt this structure, it should be reflected in their content – all sections are necessary parts of a complete and systematic scientific study!

The content also determines the order of the subsections within these sections. The section on "Data and Methods" for example can be split up into two subsections, ("Data", "Methods").

1. Introduction
- Description and discussion of the research gap or linguistic phenomenon at the hand of relevant secondary sources (research question); working definitions
- Hypothesis

2. Data and Methods
- Subjects (if applicable), e.g. participants in experiments, speakers in elicitation or corpora…
- Methodological basics, e.g. statistics, corpus or database compilation and queries, elicitation methods…
- Research process, e.g. of the experiment, of the field work
- Data, e.g. corpora, questionnaires or surveys, databases, stimuli…
- Annotations (if applicable), e.g. coding schemes, re-coding, reliability

3. Results
- Descriptive results (e.g. elicitation results) and/or descriptive and explorative statistics; representation in Figures and Tables
- Inferential statistics (if applicable)

4. Discussion
- Interpretation of results with regard to the research hypothesis
- Possible generalizations, embedding into the wider research context
- New questions that arose from the study

5. Conclusion
- Summary and outlook

In addition, the following sections are part of every research paper:

- **Table of Contents**: required for all papers
- **Table of Abbreviations**: recommended especially for papers working with glossing in examples (see Section 5.2)
- **References**: required for all papers (see Section 4.4)

Depending on its length, the paper may be divided into further **numbered sub-sections**. Too many levels of sub-sections should however be avoided – instead, additional higher-level sections provide more clarity. If sub-sections are used within one section, there need to be at least two or else the additional level is obsolete. Moreover, paragraphs without sub-heading should be avoided in sections that contain sub-sections (with the exception of short introductory paragraphs at the beginning).

There are no numbered sections below the third level (i.e. 3.1.2 is possible, but 3.1.2.2 is not). Appendices and References after the Conclusion do not have section numbering, but Appendices can be numbered as Appendix 1, Appendix 2 … etc.

At times, it makes sense to refer to passages within the paper itself. In this case, use *see* and *cf.* ('confer') in combination with section numbering, e.g. *see Section 2.4*. Note that there is a capitalized initial letter in the word *Section* whenever it refers to a specific part of the paper. When referring to passages below the uppermost level (chapter or section, depending on the text type), either the word *Section* or the paragraph symbol § may be used (e.g. §2.4).

3.2 Language

The author is free to write the paper in any of the official languages of Switzerland or English, as understood and agreed upon by the supervisor. The two most common choices are German and English.

Code-mixing is to be avoided – an English text with heavy German influences is as bad as a German text brimming with unnecessary Anglicisms. Some examples:

- *linguistic -> linguistisch, sprachlich, or Sprach-* (depending on the context)
- *agreement -> Kongruenz*
- *head -> Kopf*
- *Indo-European -> Indogermanisch*
- *Kabardian -> Kabardisch*

3.3 Scientific Register

Objectivity and clarity are the foundations of scientific writing. Thus:

- Use *maybe, could, it seems*, and similar hedging expressions sparingly. These expressions are legitimate wherever you think there is not enough evidence to make a strong statement. They should, however, not be used to mask weaknesses within the study or the argumentation. Try to make a clear statement as often as possible.
- Avoid verbose language. Concise statements are clearer and easier to understand.

- Avoid ambiguities by using scientific terminology. Because many linguistic terms (e.g. *Nebensatz*, *Kasus*, *Lexem*) are ambiguous in themselves, depending on their context, make sure to disambiguate by explaining how exactly you are using a certain technical term.

3.4 Typographical Conventions

As mentioned in Section 2.3, scientific texts should be typographically neutral. Special fonts are only to be used whenever they carry clear and well-established meaning:

- *Italicized* font to be used for:
  - linguistic examples (see e.g. Section 5.1)
  - titles of sources mentioned in the text body
  - references to terminology if they are being discussed (but not terminology in its regular application)
  - single words that need to be marked
- **Bold** font to be used for:
  - single words within linguistic examples that need to be marked
- SMALLCAPS to be used for:
  - central terms at their first mention (optional)
  - single words within linguistic examples that are being contrasted and/or marked
  - abbreviations of grammatical elements in glossing (alternatively: regular capitalization)

Avoid adding special fonts to quotation; whenever marking/italicizing still seems necessary, it needs to be acknowledged in square brackets ([emphasis mine]).

Quotation marks:

- "Double quotes" to be used for:
  - quotations (see Section 4.3)
  - terms the author wants to distance themselves from
- 'Single quotes' to be used for:
  - the meaning of linguistic examples (also in running text)

Please note that different languages have different conventions for quotation marks (e.g. German texts use either „these“ or «these» quotation marks, the latter being originally French and common in Switzerland, while the 'English' quotes mark meaning).

4 Working with secondary sources

Working with secondary sources is one of the most important learning objectives of your studies. Please read this section carefully – it contains information concerning the legal ramifications of misquoting and plagiarism as well as guidance on the topic of correct formatting of in-text citations and reference sections.

4.1 Basics: Intellectual Property

Every research paper builds on previous work. In your paper, too, you need to consider what other authors have already written concerning your topic. Your own thoughts and findings will thus stand alongside the thoughts and findings of other authors.

In principle, the contents of any scientific text will be presented as conceived by the author him- or herself. This can be based on specific data and methods, or general considerations. On the other hand, all content that was adopted from other authors has to be marked as a citation (cf. Section 4.3 below).

If this is not done properly, you are presenting another individual's intellectual property as your own. This incorrect presentation is called **plagiarism** and is not only not fair to the unacknowledged authors, it is also a violation of authorship attribution. Term papers that contain plagiarism will be graded as not successful. Depending on its severity, plagiarism might lead to legal consequences, going as far as suspension from one's studies at the University of Zurich. Please find more information about plagiarism on https://www.education.uzh.ch/en/kommissionen/lehrkommission0.html .

4.2 Which Texts to Cite

Not all texts are equally well-suited to be cited. The following factors can play a role:

- **Text type:** Encyclopaedias (including Wikipedia) and course books can be used to gain an overview of the topic, or to find secondary sources. Dictionaries can be important sources for some kinds of studies, as are descriptive grammars. Neither of those contain primary research, though, and are thus not suitable for citing. Handbooks and review articles on the other hand are suitable sources. Although they do not contain primary research, they summarize and evaluate the current state of research and its contributions to theory, and thus make their own relevant contribution.

- **Permanence.** One of the most important criteria of scientific sources is that they are permanently accessible in an unchanged form. It is for this reason that online resources are to be viewed with caution. Although there are more and more permanent resources in this area (e.g. articles in electronic research journals), many websites cannot guarantee permanence and are thus not suitable sources to be cited. You should only use resources that have been published in an established online journal and/or carry a Digital Object Identifier (DOI).
- **Age.** In principle, newer texts should be preferred to older ones. Older texts could contain results that have been falsified in the meantime. Moreover, you save a lot of time if you start with newer texts, because these usually cite the older ones and can give you an overview of the entire literature on the topic. In certain cases, however, it is necessary to also read and cite from older texts, e.g. for frequently cited classic works, or sources that have not lost their timeliness, but were overlooked by later scholarship.

A lot of insights and discoveries in research are replicated. In these cases, one should try to cite the oldest source as well as a small but representative selection of younger sources, e.g.

> Grammatical relations are construction specific (Van Valin 1981; Dryer 1997; Croft 2003; Bickel 2011).

4.3 In-text Citing

When citing contents from another source, you have to mark them as a citation. You can choose to quote directly from the text (literal quote) or cite indirectly.

- Short literal quotes have to be put into double quotation marks, e.g. "Colorless green ideas sleep furiously".
- Literal quotations of more than one line in length should be spaced out and indented. Quotation marks are hereby optional. For example:

  > "Eine Nation hat freilich im Ganzen dieselbe Sprache, allein schon nicht alle Einzelnen in ihr, wie wir gleich im Folgenden sehen werden, ganz dieselbe, und geht man noch weiter in das Feinste über, so besitzt wirklich jeder Mensch seine eigne."

- For indirect quotations, please ensure to be transparent as to where the citation starts and where it ends, e.g. by adding the source in brackets at the beginning and/or end of the account (Greenberg 1962), or by using phrases like *According to Greenberg (1962),…* or *Greenberg (1962) distinguishes…*

In general, indirect quotes are preferred over literal quotations, as the exact wording of the author is almost never of crucial importance in linguistics (as opposed to e.g. literature studies or philosophy). Moreover, frequent literal quotations negatively impact the readability of the text.

Each citation needs to be clearly attributed to the source it is taken from. In linguistics, the conventions are to refer to author and date of the source. This pair then refers to a detailed bibliographic entry in the References section at the end of the paper (see Section 4.4 below). For literal quotations, the page from which the passage was taken has to be added to the citation. Whenever an indirect quotation refers to a specific passage in the source, the page(s) should also be added to the citation. For best practice examples, see the following:

- Usually, author and date are put into brackets, e.g.

  > It has been claimed repeatedly that colorless green ideas could sleep (Chomsky 1957).

- Whenever the author is spelled out in the text body, just the year stands in brackets:

  > Even Chomsky (1957) was concerned with the furious sleeping of ideas.

- Page(s) are separated by a colon following the year:

  > "Colorless green ideas sleep furiously" (Chomsky 1957:15)

- A couple of established abbreviations can be used to refer to pages (but they are considered rather outdated and have gone out of use with a number of publishing houses):
  - *f.* (Ger. *folgende* 'the following', originally Lat. *folio*) refers to two successive pages; *ff.* (Ger. *fortfolgende*, originally Lat. *folio et foliis*) refers to more than two successive pages. Instead of these, the exact range of pages (e.g. 67–72) should be spelled out.
  - *op. cit.* (Lat. *opus citatum* 'the cited work') refers to the source cited last, without information on the page.
  - *ibid*. (Lat. *ibidem* 'in the same place') refers to the exact same source and page cited last.

- With co-authored sources, both have to be named in the text, linked by an ampersand (&). For more than two authors, use the first author's name and the abbreviation *et al.* (Lat. *et alii* 'and others'). In the Reference section, all authors have to be listed. Examples:

  > Gundel et al. (1993) relate definiteness to a givenness hierarchy, which incorporates the common concepts of familiarity and uniqueness. However, neither of these is truly sufficient for defining definiteness (Birner & Ward 1994).

- When citing several sources by the same author, the years are separated by commas. Whenever sources by different authors are cited in brackets, the author-date units are separated by semicolons:

  > Lasagnefüllung kann mit gehackten, gekochten Eiern verbessert werden (Tommasi 2002, 2005); besonders Secondo-Chefs bevorzugen aber die Variante ohne Ei (Forchetta 1989; Cucchiaio 2012, 2014; Coltello 2014).

- Whenever the Reference section contains several entries by the same author published in the same year, author-date reference is ambiguous. Add small letters a, b, c, … after the year to avoid ambiguity:

  > P-labile verbs can be classified according to degree of transitivity (Letuchij 2009a) and to the relation holding between the transitive and the intransitive use (Letuchij 2009b).

- Usually, direct quotations are cited verbatim. If they are changed for reasons of readability, omissions are marked by […] and additions by [text, author's note], for example:

  > So bemerkt Gspassettl (1942:8), gerade in der heutigen Zeit würden «Randen, Krautstiel und Kefen [...], ehemals Embleme der Znüni-Kultur, immer häufiger zum Zmorge [i.e. breakfast, author's note] verspeist».

- If a quotation contains an obvious typographic mistake and you would like to clarify that it was part of the original quote, you can add [sic!] (Lat. 'thus'):

  > Gspassetl (1942:9) fährt fort: "Kohlrobis [sic!] sind hingegen weniger beliebt."

- Other than citing sources, you can alert the reader to relevant sources by using *cf.* 'confer'. This should be done sparingly, however, and certainly not for reasons of increasing the number of cited works.

- The title of a cited source usually only appears in the Reference section. If it does occur in the text body, it should be put into italicized font.

4.4 The Reference Section

The reference section contains detailed information on all sources cited in the text. It is not customary in linguistics to include sources that are not cited in the text. Citations and References are linked by the author-date units in the text body, including added letters to avoid ambiguity, if necessary (see Section 4.3 above).

Each entry in the Reference section needs to contain the following information:

- Author(s). The References have to be listed alphabetically. Editors are marked as *ed.* (sing.) or *eds.* (pl.).
- Year of publication. In the case of pre-publication sources, use either:
    - *in preparation*
    - *under review* or
    - *in press*
- Title of the source

All additional information depends on the type of publication:

- **Book/monograph**: publisher and place of publication.
- **Journal and Proceedings articles**: name of the journal or proceedings title, volume – plus issue number where applicable, e.g. 43(3) – and exact page range.
- **Article in an edited book**: editor(s), title of edited book, publisher and place of publication, exact page range.
- **Presentation**: Scientific presentations or talks may be cited in cases in which the information is not (yet) available in written/published form. In these rare cases, the context of the talk (e.g. conference) as well as time and place need to be listed. If there are additional materials that the text refers to, such as a handout or presentation slides, these need to be made accessible (e.g. by linking to a website).
- **Websites:** Regular websites are not suitable to be cited in a research paper (see criterion of permanence, Section 4.2). Text published on Online platforms however can be cited, following the same guidelines as paper publications. Whenever possible, the DOI (Digital Object Identifier) has to be

added. Hybrid publication types that are not one of the now established publication types can only be cited with DOI.

There are multiple, sometimes contradicting, style sheets concerning the formatting of this information – they all pay tribute to different aesthetic traditions. Above all, consistency in formatting is key, more so than the choice of a particular citation style. It is, for example, less important whether the year is put in brackets or not, and more important that the style is followed consistently throughout.

The most straightforward way of achieving consistency is by using a formatting template. The Department of Comparative Language Science recommends the *Unified Linguistic Stylesheet* (ULS, https://www.linguisticsociety.org/resource/unified-style-sheet), which represents the results of a joint effort to combine the traditions of various linguistic journals.

In the following, please find the stylistic conventions of the ULS. The *Committee of Editors of Linguistic Journals* provides a Bibstyle-template for LaTeX, which can be found in a slightly updated version (which keeps capitals in titles intact) on the homepage of the DCLS.

- Information units (e.g. author, date, …) are separated by full stops/periods within each entry.
- The first author is listed first in the form "last name, first name"; additional authors follow as "first name last name"
- This is followed by the year of publication.
- Titles of independent publications are put into italics, and titles of dependent publications (articles, contributions to edited volumes, …) in regular font.
- Capitalization only applies to English names of publishers, journals and book series, where all content words are capitalized. Everywhere else, standard spelling and capitalization applies.

The end of the reference entry also depends on the text type:

- **Book/monograph**: "Place of Publication: Publisher."
- **Journal article**: "*Journal Name* volume. page range."
- **Contribution to edited volume**: "In editor first name last name (ed.) *Title*, page range. Place of publication: Publisher."

Please see the exemplary references below.

Abbott, Barbara. 2010. *Reference*. Oxford: Oxford University Press.

Bach, Kent. 2006. The Top 10 Misconceptions about Implicature. In Betty Birner & Gregory Ward (eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, 21-30. Amsterdam/Philadelphia: John Benjamins.

Ebert, Karen. 1997a. *Camling*. München: Lincom.

Ebert, Karen. 1997b. *A grammar of Athpare*. München: Lincom.

Gundel, Jeanette, Nancy Hedberg & Ron Zacharski. 1993. Cognitive Status and the Form of Referring Expressions in Discourse. *Language* 69(2). 274-307.

Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535-689.

Kifle, Nazareth Amlesom. 2007. Differential object marking and topicality in Tigrinya. In Miriam Butt & Tracey Holloway King (eds.), *Proceedings of the LFG07 Conference*, 5-25. Stanford: CSLI.

Lamsāla, Rāmcandra. 2062 V.S. (= 2005/2006 AD). *Nepālī bhāṣā ra vyākaraṇa*. Kirtipur: Sunlight publication.

Mohanan, Tara. 1994. *Argument Structure in Hindi*. Stanford: CSLI.

Peirce, Charles. 1906. Prolegomena to an Apology for Pragmaticism. *The Monist* 16(4). 492--546.

Sprigg, Richard Keith. 1989. Review of A Grammar of Limbu by George van Driem. *Bulletin of the School of Oriental and African Studies* 52(1).

Winford, Donald. 2010. Contact and Borrowing. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 170-187. Chichester: Wiley-Blackwell.

5 Linguistic Conventions

5.1 Transcription

Linguistic examples have to be put into italicized font, also in the first line of glossing examples (see Section 5.2). Exempt from this rule are phonetic-phonological transcriptions, which can be marked through the use of [square brackets] (for phonetic) or /forward slashes/ (for phonological examples).

Transcriptions based on the Latin script are preferred whenever possible; they should represent the pronunciation directly and unambiguously. A very widespread universal transcription standard is the International Phonetic Alphabet (IPA, https://www.internationalphoneticassociation.org/content/full-ipa-chart).

At times, the decision on a transcription system can be influenced by a number of language-related and / or philological traditions – or by practical considerations. Here are some frequent exceptions:

- If a language is customarily transcribed in **Latin script** in scientific publications, this should be chosen, apart from passages in which phonetic-phonological information is needed. In the latter case, IPA is the best choice. Keeping to orthographical conventions also improves the comparability across text corpora and publications. This also applies to languages that have long established spelling conventions, such as English, Vietnamese, or French, even if they are phonologically non-transparent.
- If a language has spelling conventions that are traditionally represented in **non-Latin scrip**, a transcription must be used. This also applies to alphabets closely related to the Latin one, such as Greek or Cyrillic. Many languages have philological traditions for these transcriptions, which should be referred to. IPA should only be used where no accepted transcription exists or where a transcription is considered unfit.
  An additional use of the original script is usually only justified if this is needed for the argumentation, or whenever it carries additional relevant information (e.g. in graphological studies). An exception to this is presented by Sinitic examples, in which the original script provides important information for lexeme identification to the reader.
- For languages **without orthography**, the IPA is the standard transcription system. If there is a repeated occurrence of single sounds for which the IPA system is cumbersome, they may be represented in a different manner. For example, many Nepalese languages contain the phoneme /t̪s̪/, which is spelled ⟨c⟩ according to the tradition of transliteration followed in the field of Indology. It is important to clearly spell out in those cases which sounds are not being represented according to the IPA (for example, many authors prefer y over j for the palatal glide, even if the rest of the transcription is in IPA – this must be clearly stated).

Apart from the transcription, consistency is key. Transcription systems should never be mixed within one text. The only exception to this is the IPA, which can always be used alongside another transcription system, whenever needed for phonetic-phonological precision.

There are no guidelines as to how broad the IPA transcription needs to be. Usually, broad phonological transcription is sufficient; this has to be decided for each individual case. The determining factor will always be whether – and if yes, which – phonological findings are relevant for this particular paper. In syntactic papers, for example, phonological detail is usually not particularly relevant.

Quotations and citations of sources that are written in non-Latin script are a special case:

- Here, indirect quotations should take preference to an even greater degree. If a direct quote in a language other than English (or the language of the paper) is still necessary, a translation has to be submitted (see Section 5.2 below). In this case, either transcription or original script can be used, because the meaning becomes apparent from the transcription and the pronunciation of the original quotation is not relevant.
- For Reference entries, it is crucial that each source can be clearly identified. Transcriptions that are phonologically true but do not represent the spelling are thus unsuitable. The Reference section thus has to contain either the original script or an exact transliteration. For in-text citations (when referring to the author's name), transliteration should be used for ease of reading.

5.2 Glossing and Translations

All linguistic examples have to be translated. Translations are to be put into 'single quotation' marks.

Linguistic examples that consist of more than one word or noun phrase also have to be glossed. This entails segmentation and annotation morpheme by morpheme. Glossing is also recommended for single words if they consist of morphologically complex word forms. The standard for glossing in use at the department is the *Leipzig Glossing Rules* (http://www.eva.mpg.de/lingua/resources/glossing-rules.php). The metalanguage for glosses is English, independent of the language in which the paper is written.

Longer linguistic examples are to be indented and numbered. They contain, apart from to the glossing and translation, a source that refers to either a published/cited work or to the author's own corpus, such as:

(1) *Hani-bakhra-ce sa-ŋa ghãsa pid-u-c-e?*
2pPOSS-goat-ns who-ERG grass give-3O-3nsO-IND.PST[.3sA]
'Who gave grass to your goats?' (CLC:CLLDCh1R07S03.068)

Glossing often contains abbreviations of grammatical morphemes. These abbreviations are usually put into capitals or smallcaps. The *Leipzig Glossing Rules* contain a list of recommended abbreviations for the most common grammatical functions. Additional abbreviations that go beyond this list can be introduced but have to be explained in a Table of Abbreviations (see Section 3.1).

Translations may contain elements that are not contained in the linguistic example, but added from the context; these are put into square brackets, for example 'they hit [him]' (when *they hit* is ungrammatical in English), or 'they hit him [the gardener]' (when *him* would be ambiguous without context). The translation should aim to represent the semantic content of the whole example instead of translating word by word. If these two meanings are far removed, a more literal translation can be added in brackets, e.g. 'they did see him (lit. their eyes got caught by his figure)'. Linguistic examples can be subdivided:

(2) a. Kina asu ghanta rok-e a-lis-i-hẽ na?
SEQ how.many hour stop-V.NTVZ 2S-be-p-IND.PST CTOP
'So for how many hours did you stop?' (CLC:dkt-ktm-trip.0264a)

b. Pahiro rok-e num-ma jamma boll-a numd-i-ne.
landslide stop-V.NTVZ do-INF all effort-NTVZ do-1pi[S]-OPT
'We should all make an effort in order to stop the landslide.'
(CLC:ctn_prob_talk.0719)

c. Gadi rok-e mett-u-ku-ŋ.
car stop-V.NTVZ do.to-3[s]O-IND.NPST-1sA
'I stop the car.' (CLC:tangkera_04.479)

5.3 Terminology and Frameworks

The Comparative Linguistics study program does not require the use of special terminology. Established terms should be used whenever possible; if multiple terms coexist, one has to be chosen and used consistently.

It is crucial that the intended meaning of terminology becomes transparent. In the field of linguistics, a scientific term can have different meanings depending on the theoretical background in which it is being used. Whenever ambiguities run the risk of resulting in misunderstandings, scientific terms that are of central relevance to the paper should be defined (e.g. in the Introduction or at the beginning of the main section, see Section 3.1).

Along the same lines, the study program does not provide the theoretical framework for a paper. If you are insecure in matters of theory, we recommend you follow a framework that was introduced as part of your studies; otherwise you are free to choose one. Again, internal consistency should be a guiding principle.