



MANUAL

For Papers in English Linguistics

linguistics noun [ling-gwɪstɪks]
: the science of language, including
morphology, syntax, semantics,
and historical linguistics.
— *Merriam-Webster's Collegiate Dictionary*, 11th ed., 2003

**Gunnel Tottie et al.
University of Zurich, English Department
1999/2008/2012/2017/2023**

Fourth revised version (2023) by Omar Itani, Olga Timofeeva, Simone E. Pfenninger and Marianne Hundt.

Third revised version (2017) by Larssyn Staley, Adina Staicov, Nicole Eberle, Daniela Landert and Lena Zipp.

Second revised version (2012) by Sarah Chevalier.

Revised version (2008) by Anja Janoschka and Dorota Smyk-Bhattacharjee.

With many thanks to Jan Svartvik, Martin Heusser, Britt Erman and Miriam Locher (1999), Andreas Fischer, Andreas H. Jucker and Barb Breustedt (2008), Daniela Landert, Andreas H. Jucker and Danielle Hickey (2012).

Cover images: <<https://www.ilovelanguages.com/best-ways-to-learn-english/>> and
<<https://www.bath.ac.uk/guides/introduction-to-english-linguistics/>>

Table of Contents

1. Introduction: Linguistics papers and literary essays.....	1
2. Research and writing	1
What are the purpose and objectives of writing a linguistics seminar paper?	1
Choosing your subject	1
Length of your paper and deadlines	2
Sources and how to use them.....	2
Doing the research	2
Plagiarism and Artificial Intelligence	3
Doing the analysis	4
Support your statements	4
Who are you writing for?	4
Ethics	5
3. How to organize your paper.....	6
Table of contents	6
1. Introduction	6
2. Previous work.....	6
3. Data or participants.....	7
4. Method.....	7
5. Results and analysis.....	7
6. (Discussion)	8
7. Conclusion	8
Sources	8
References	8
Appendix (or Appendixes/Appendices; either spelling is fine)	9
Footnotes	9
3. Tables and figures: Displaying and summarizing your data	9
Tables	9
Figures	10
Normalized figures and percentages	11
4. Statistics: Some basics	11
Descriptive versus inferential statistics.....	11
Sample size and reliability.....	12

Questions arising from a limited size sample.....	12
5. Style sheet.....	13
Introducing concepts.....	13
Emphasis	14
Quoting linguistic examples from your data	14
Quoting from the literature	14
References in the text	16
Formatting the reference section	16
Sources/Primary sources.....	17
Print references/Secondary sources	18
Online references.....	20
6. Checklist for papers	21
References	22

1. Introduction: Linguistics papers and literary essays

At the English department of the University of Zurich, we teach both English Linguistics and English Literature, as well as Language Skills and Culture. The two principal disciplines have different research traditions, and scholarly publications in all three fields have different rules for organizing and formatting papers. This booklet tells you how to carry out research and organize your data for a linguistics paper as well as how to quote and give references according to the practices of linguists. Please note that these practices are different from those of literary and critical writing as detailed in Kavka/Schweizer et al. (2022). However, the rules for writing good English are the same for literary and linguistics scholars, and there are many useful tips in the Kavka/Schweizer et al. booklet (<https://www.es.uzh.ch/en/studies/studyadvice/furtherinformation/downloads.html> > English Literature > Literature Guide).

2. Research and writing

What are the purpose and objectives of writing a linguistics seminar paper?

A seminar paper is an exercise in linguistic research. It is a process of acquiring new knowledge based on formulating research questions/hypotheses and collecting measurable evidence. You have to show that you can do the following:

- Search for and identify relevant research;
- Identify a research gap in a specific field;
- Formulate a research question that aims to fill this gap;
- Select and apply the appropriate theory, method and data to answer your research question(s);
- Limit the scope of your discussion and analysis;
- Conduct a theoretical discussion and analyze linguistic data in English;
- Critically assess previous research and your own data set.

Choosing your subject

It is important to make a realistic choice of topic and scope and to pick something that can be completed in time. The appropriate scope of your project will depend on whether you are working on an *Introduction to Linguistics* term paper, a BA seminar paper or thesis or a MA seminar paper or thesis. You may be inspired by a research paper you have read and test their theory on your own data or compare their results to data you have collected. You may also notice a linguistic phenomenon, such as (1) *your guys'* being used as a plural genitive, (2) the frequent use of a particular intensifier (e.g. *super*), or (3) *be like* being used as a quotative marker (e.g. *and I was like*), and investigate who uses this form and in what context.

Often you will propose your own research topic. You are, however, strongly advised to discuss your ideas with your instructor to check relevance, scope, methodology and feasibility. Note that it is your responsibility to take the initiative and that the amount of freedom/help will depend upon the course and instructor. Your instructor may also assign the task of writing a paper proposal.

Length of your paper and deadlines

- Introduction to Linguistics papers: 3,500–4,000 words
- Bachelor research papers: 4,000–5,000 words
- Bachelor thesis: 8,000 words; 1 December (HS) / 1 June (FS)
- Master research seminar: 8,000 words
- Master thesis: 20,000–30,000 words; 1 December (HS) / 1 June (FS)

The word count excludes references and appendices, but includes tables, figures and examples.
Seminar papers have individual submission deadlines.

Sources and how to use them

Sources consist of data sources (or primary sources) and relevant literature (or secondary literature). Data sources are the actual linguistic data that you analyze and write about, such as corpus and dictionary data, experimental data, transcriptions of spoken material (e.g. recordings and/or elicited responses from speakers). The relevant secondary literature is what has been previously written on the linguistic topic you have chosen.

Doing the research

There are multiple starting points for research in linguistics. You may start with a research gap you found in a research paper, by applying a theory to a different data set, or with a linguistic feature that you find interesting. Once you have a general topic, you need to further specify it by reading previous literature dealing with various aspects of your topic. Based on what has already been done, you will better define your topic, refine your research questions and select the best data set and method to answer your research question(s).

Use specialized dictionaries of Linguistic terminology (by Brown & Miller, Crystal or Matthews, see references) rather than the OED to define the linguistic feature you are studying.

For a systematic search of previous research, the best starting place is Brill's Linguistics Bibliography online: <https://brill.com/display/db/lbo>

The ES library website provides a detailed list of catalogues and electronic resources. The following list only includes links to the most important of these resources with some comments on how to use them.

University Library of Zurich: <https://www.ub.uzh.ch/>

Swisscovery database: <https://swisscovery.slsp.ch/>

Information about books and journals available at UZH libraries and other Swiss libraries.

Note: Make sure to go to tab “with CDI” (in the German interface: “mit CDI”) if you want to search for literature including journal articles and book chapters. If you use the first/default search tab, your search will only return complete books and you will miss much relevant research.

Elektronische Zeitschriftenbibliothek (EZB): <https://ezb.ur.de/ezeit/index.phtml>

Overview of all relevant journals, information about whether/how they can be accessed (i.e. whether you have online access through the ZB), and links to online access.

Note: You need to access the journal pages from within the University network or through VPN to have access to journals to which the ZB subscribes (yellow dot). To configure VPN, see <https://www.zi.uzh.ch/en/students/workplace-collaboration/vpn.html>.

For a small fee, you can order books from other libraries to the ZB or to your home. Finally, **JSTOR** (<https://www.jstor.org/>) and **Google Scholar** (<https://scholar.google.com/>) are viable resources, too.

You may also consult your supervisor for references. When reading previous literature, make note of key findings and the method used. An awareness of previous findings will help you formulate hypotheses and research questions for your own study and situate your work in the field. Additionally, by paying attention to the methods used, you can make more educated decisions about how you should conduct your research. When taking notes, make sure that you always include bibliographic information, i.e. the reference (see below), and that you differentiate between direct quotes and paraphrases in your notes, so that you can correctly cite the sources in your own paper. You may find it useful to use a database like **Endnote** or **Zotero** to manage your references and notes. Copy your excerpts correctly and note the name of the source and the page number immediately. This will save you a lot of extra work and trouble later on.

Once you have a clear idea of what you want to research and what factors you should consider, it is time to start collecting and analyzing your data. Your own data analysis may require you to read and include more previous research on a specific area of your topic. Research is a circular process: your initial reading will provide a foundation and help you set up your study, but once you have started your own analysis, you may often find that you need to read previous research in a more specific area. During this process, you may also find your research question becoming more and more focused.

Plagiarism and Artificial Intelligence

According to Wray and Bloomer (2006: 237), “[p]lagiarism is the theft of other people’s words and ideas. Plagiarism happens when you claim (or appear to claim) that an idea, or the expression of it, is your own when in fact it is someone else’s”; it is, therefore, an infringement of copyright law. The rule of thumb is that any idea which is not your own must be referenced. The only exceptions are facts which can be considered general knowledge, e.g. “Saussure was a Swiss linguist”. If you are unsure about what can be considered general knowledge, it is better to over-reference. Avoiding plagiarism starts at the research and note-taking stage. Make sure that you are reading enough previous work to include a variety of perspectives, and when taking notes, always include the source and enclose direct quotations in quotation marks.

Similarly, avoid autoplagiarism. You must not submit work for assessment which you have already submitted (partially or in full) to fulfil the requirements of another degree course or examination.

(Auto)plagiarism can result in suspension from seminars, lectures, and one’s entire studies. For more information on plagiarism, see English Department guidelines at <https://www.es.uzh.ch/en/studies/studyadvice/furtherinformation/downloads.html> > General Guidelines and Templates > *What constitutes Plagiarism?*).

Similar principles apply to the use of Artificial Intelligence (AI). You must cite and reference any text or output generated by AI in your assignments. Remember that it is a tool to assist and inform you, but not to replace for your critical thinking and analysis. For more information on good practice when using AI, see English Department homepage > General Guidelines and Templates > *Guidelines on ChatGPT and AI*).

All papers need to be submitted with a signed statement of authorship (see the same homepage > General Guidelines and Templates > *Selbstständigkeitserklärung*). Additionally, all papers must be submitted in electronic form and will be checked using *Turnitin*, a plagiarism detection software.

Doing the analysis

If you have conducted a quantitative study of a linguistic phenomenon, it is often best to present your data in tables and/or figures. A good way to start is to draw up your tables first, and then to go on to describe key findings of the data presented in the tables and figures. Note that you cannot just provide data in tables; you need to describe what you see in the tables and how you interpret it in relation to your research question(s)/hypothesis and previous research. Additionally, the data presented in your tables and figures should be further substantiated through examples from your data and analyses of those (for more details, see Chapter 4). When presenting your results, you should always return to your research question(s) and focus on the data that provide an answer to your specific question. On how to structure your study and where analysis belongs among other sections of your paper, see Chapter 3 below.

Support your statements

In order to answer your research question, you will have to make claims about your data and findings. When making claims, refer to what can be seen in your own data and to how your findings and/or claims are similar or different from previous research on the topic. If you quote previous literature, include as much as is necessary, but not more. Often it is more appropriate to paraphrase something than to quote directly, but in either case the original author needs to be cited.

The strength of your claims should equal the strength of your findings. In other words, if you use *significant* as in “there is a significant difference in the use of *like* between the age groups studied”, then such a statement needs to be supported with results from a statistical test (e.g. chi-square or log-likelihood) that show that the difference in the results is significant enough, i.e. that they are not due to chance. Statistical testing is discussed in more detail in Chapter 5 below.

Modal verbs, such as *may*, *might*, *could*, adjectives and adverbs, such as *probable* and *possible*, and additional hedging, such as the *data suggest/indicate*, may be needed when making claims based on your findings. You want to avoid anecdotes, be as specific as possible, and provide examples from your data to support your claims.

Who are you writing for?

Academic writing is formal; it should be detailed and precise, but it should not be unnecessarily complicated. The key to good academic writing is to be able to present complicated matters in

an understandable and accessible way. It is important to have the intended reader in mind when writing. Your reader is educated in linguistics, so you should use technical terms. However, different sub-fields of linguistics and specific research groups sometimes use terms and define concepts in different ways. Therefore, it is necessary for you to present how you define the relevant concepts in your study (for instance, in a theoretical background section: see below). As a rule of thumb, think about fellow students as intended readers. How much information do you need to provide so that they are able to follow your paper.

Ethics

Linguistic research involves human informants (also referred to as participants, informants or collaborators). They produce the language, written or oral, that is studied. It is of the utmost importance that the participants provide “informed consent”.

Informed consent: According to the American Psychological Association (APA), participants need to be informed about the purpose of the research, the expectations of them and the procedures. Participants need to be informed about how the data they provide will be used and what will happen to their data after the present project: i.e. in what way will it be published, and will it only be used for the present project or will it be used for future projects? Additionally, participants need to be aware of their right to decline to participate and to withdraw from participation and of any limitations there might be due to confidentiality. Finally, participants should also be informed about who to contact when they have questions about the research (APA 2017). If you recruit minors, signed parental consent needs to be obtained. That said, children and young people should be involved in the decision-making process whenever possible. You should ensure that they receive information about your study that is understandable to them.

Contact your supervisor to discuss suitable consent sheets.

Confidentiality and anonymity: D’Arcy and Young (2012) discuss how merely changing a participant’s name is often insufficient. In some cases, the content of the data may point towards a specific group of people, even if specific individuals cannot be identified. The researcher and the participants need to be aware of what effects this might have.

Social media as a source: As D’Arcy and Young (2012: 535) also point out, social media blur the boundary between public and private. Even though social media users have chosen to make something public, or public within their network, this does not suffice as consent to their posts being used for research purposes. E.g. “[f]acebook policy allows the collection of information by third parties from its site, but users’ consent must be obtained” (D’Arcy and Young 2012: 536). As a user of social media, researchers may have access to their own network and extended networks through comments made by network outsiders in reference to people within one’s network, yet care needs to be taken to receive consent from all parties involved in the discourse being studied.

Audio/video recordings and questionnaires as a source: As researchers of linguistics, we aim to record language without affecting the production of language and get around the “observer’s paradox” (Labov 1972: 209). However, as stated by the British Association for Applied Linguistics (hereafter BAAL: 2010), “[d]eception is unacceptable because it violates the

principles of informed consent and the right to privacy". So that participants are not focused on the specific language feature being investigated, one could set up multiple tasks for the participants, to slightly distract them; this works well in questionnaires or discourse completion tasks. In interviews, BAAL suggests only informing the participants of the general aims rather than the specific aims. For more information on conducting fieldwork, see Schreier (2013).

3. How to organize your paper

A seminar paper should normally contain the sections listed below. They should be numbered, except for the Table of contents, the References, and the Appendix. Note that the section titles which follow are generic and may be adapted to your topic, e.g. "Previous work on...". Note also that the word *section* is appropriate for seminar papers; *chapter* is only used for longer pieces, such as a PhD dissertation.

Table of contents

Indicate sections and page references. Page numbering starts on the first page of the text, not with the Table of contents. Indent subsections for clarity.

1. Introduction

A good introduction starts by presenting the focus of the paper, such as "The present paper focuses on the use of plural terms of address/intensifiers/modal verbs in Canadian English". This should be followed by research-based background information on the topic, which should then lead to your specific aims and scope (including a reflection on possible limitations), research question(s) and/or hypothesis/es and a general statement about how you intend to answer your research question(s) (i.e. a corpus-based approach; a quantitative/qualitative analysis of modal verbs in the sitcom X, etc.). The background information that you provide on your topic should focus on the linguistic feature that you are studying, so if you are studying the use of intensifiers in a particular sitcom, you should focus on research on intensifiers, rather than on the sitcom itself. Information on the linguistic area that you are studying should also lead to a statement about your motivation for the specific research question. Motivation for a certain project should stem from a research gap. Note that in linguistic research papers it is not necessary to provide a number of hypotheses. Papers in linguistics should be structured around a research question and the aim should be to provide a convincing answer to the question based on your observations.

Your introduction should move from the general topic to your specific aim, in the shape of a funnel. Your introduction should delimit the scope of your study and tell the reader what aspects you intend to investigate as well as what will be left out. Your introduction should end with a paragraph providing a roadmap of how your paper is organized, i.e. "In the following section, I will..." As a rule of thumb for seminar papers, your introduction will comprise up to 20% of your paper.

2. Previous work

This section contains a critical survey of earlier work dealing with your subject. The goal is not to provide a list of summaries of previous research on the general topic, but rather to synthesize previous work to form a foundation and basis for how and why you are carrying out your

research with your chosen research question, theory and method. You should draw connections between the studies, such as how their results (key findings) build on or contradict each other and how their methods provide different perspectives on the topic. You also need to make explicit connections between previous research and your research, i.e. describe which aspects of previous studies are important or relevant for your approach. Your study should be influenced by previous work, and this may lead to a statement of a more specific research question or aim at the end of the previous work section.

3. Data or participants

In order to answer your research question, you will need a certain type of data (or ‘material’). In this section, you should introduce your data / your participants. If you are using a corpus, state which one and the specific section(s) that you are and are not using. If you are collecting your data through interviews or a survey, who are your informants? How did you devise the interviews/survey? What procedures you followed? If you are using a television show or film, where did you get the transcripts and how did you verify how authentic they are to the actual show or film? If you are working with editions of historical manuscripts, according to what principles were they edited?

Additionally, the size of your data set, as well as socio-psychological and demographic factors (e.g. age, gender, socio-economic group, IQ, cognitive functioning, language history) that you have accounted for should be outlined, and your variable should be defined along with a description of what you are including and excluding (e.g. for the study of intensifiers, are you only looking at positive or negative intensifiers). Also, illustrate your variable with specific examples from your data set. In this section, you should also state the nature and limitations of your data. Consider how your choice of data affects the results and analysis.

4. Method

Explain your method of analysis and mention the advantages and limitations of your approach. This should include an explanation of how you identify or extract your specific variable (e.g. terms of address, intensifiers, modal verbs, etc.) from the data set, as well as how you categorize it. If you are using a typology from previous research, for example, are you using the exact original typology, or have you modified it by adding or removing categories to fit your data? How did you transcribe and code your qualitative interviews? What software did you use? What statistical tests did you employ and why?

If your investigation is long and complex, give a step-by-step description of what you did. For example, if you use complex search strings in a corpus, you should provide them for your reader. If you transcribe your data or modify existing transcriptions, outline the criteria you use. As a rule of thumb, your explanations should be sufficiently detailed that someone else can follow them and arrive at the same results.

5. Results and analysis

First you present your results, and then you analyze them. Tables, figures, and examples can all be used to present your results. After organizing your results into one or several of these forms, it is your job to guide the reader through your analysis. You need to point out your key findings

(i.e. the ones that support the answer to your research question), and you need to explain how they support the answer to your research question. In your research, you will usually also encounter negative evidence, i.e. findings that contradict your claims. You should not ignore this, but rather acknowledge and discuss it in addition to the evidence that supports your claims.

Make your analyses as clear as possible and concentrate on one aspect at a time. Support your arguments with references to your data (which you first presented in the tables, figures, and examples mentioned above). You may wish to divide your results and analysis into several sections, especially in longer papers which have multiple components in the research question. Long and complicated sections should have a short summary at the end.

6. (Discussion)

Some papers include a discussion section which pulls the key findings of the present research together and situates them in relation to previous key research findings presented in the previous research section. This helps to create a complete circle and show the relevance of the present research for the specific field of linguistics.

If you choose to not include a separate discussion section, this information should be present in either the analysis or conclusion section. Keep in mind that no new information should be presented in the conclusion.

7. Conclusion

In this section, you need to revisit the research question by providing an answer and supporting it through a summary of your key results: Which findings support your conclusion, and which might point to another conclusion? Additionally, address how your study contributes to the specific field outlined in the previous work section and what the main practical (pedagogical, social...) implications are. If parts of your results are inconclusive, e.g. because you have not had enough material, say so. Also indicate what aspects or areas demand further study in order to formulate an outlook.

Sources

If your data consist of published material (e.g. newspapers, dictionaries, films, electronic corpora, websites), these can be listed here in a separate section as “Sources” or under References (see below) as “Primary sources”.

References

Under References (also Works cited or Bibliography), you list all the publications that you refer to in your text in alphabetical order. You can divide your references into “Primary sources” (see above) and “Secondary sources”. Only those publications that are mentioned explicitly in your paper should be included. The sources that you reference should be varied; a range of journal articles, monographs, handbooks, etc. should be used. This should provide multiple levels of insight and perspective. Additionally, cited literature should include up-to-date research and be relevant to your specific project.

Appendix (or Appendixes/Appendices; either spelling is fine)

In linguistics, appendices are used to present a questionnaire used to collect data, long data extracts, extended examples, and additional information on the data that is not essential to understanding the results and that would disturb the flow of reading. If you use an appendix, it should be mentioned in the text, but your reader should be able to read and follow the argument of your paper without having to refer to it. Appendices should have separate numbering (Appendix A, Appendix B, etc.) but continuous page numbers with the rest of the paper.

Footnotes

Footnotes are also used sparingly for additional information which you feel might interrupt the flow of your argument. However, the same argument as above applies: your reader should be able to read and follow the argument of your paper without having to refer to a footnote.

3. Tables and figures: Displaying and summarizing your data

Tables

Very often a table is a good way of displaying results of a quantitative nature. Tables help the reader to grasp what your results are. For instance, if you are reporting on the occurrence of different modal verbs in literary works, you may present your data in a table (see Table 1 below). There are several aspects you should note:

- Every table needs a caption (i.e. Table number and Description);
- Numbers in tables should be adjusted to the right to ease comparison;
- When data sets vary in size (e.g. *Great Expectations* is 186,564 words and *The Importance of Being Earnest* 20,740 words), frequencies need to be normalized to enable comparison between data sets. Normalized frequency = raw frequency / token number * common base. The token number is the number of words in the respective data set; the common base corresponds to the number in “frequency per 1,000/10,000/etc. words”.

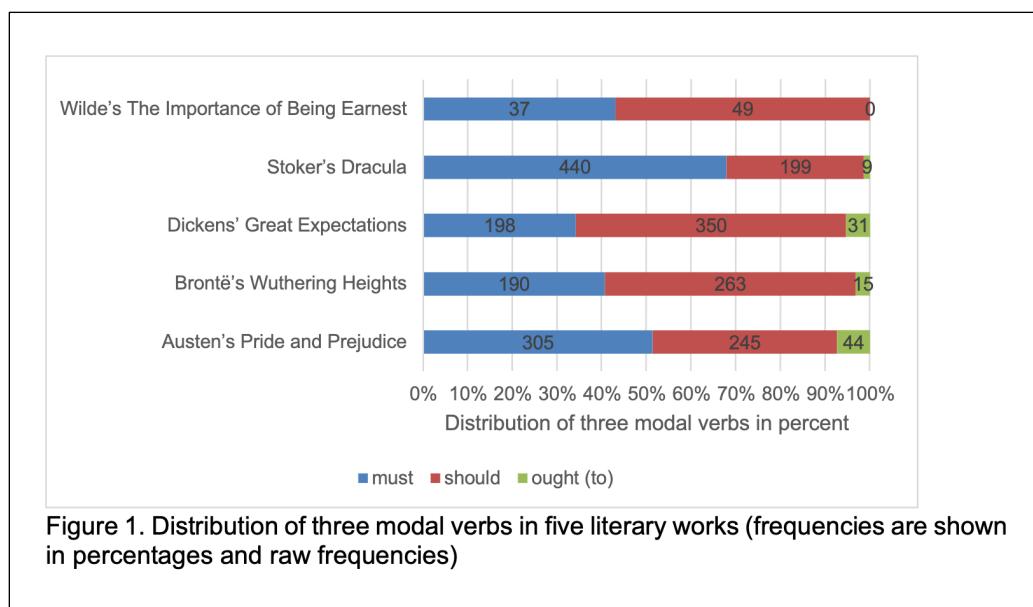
Note also that even though a table is a practical way of summarizing information, it does not free you from also describing and highlighting your key results in the main text either before or after the table or from positioning them among previous findings and/or theories. Although you can describe the data in a table before or after it, a new section should not start with a table.

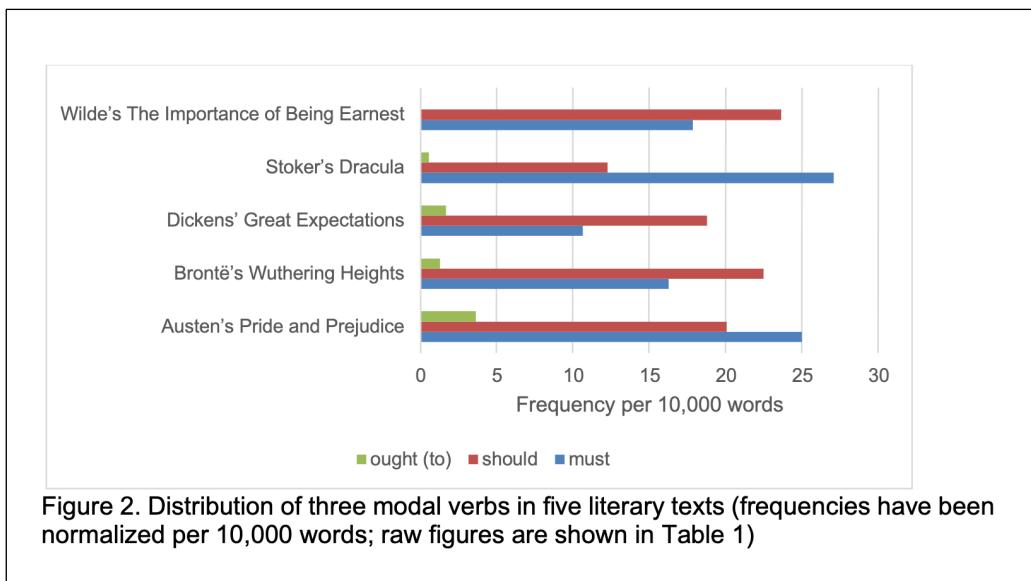
Table 1. Frequency of three modal verbs in five literary works, normalized per 10,000 words (and absolute)

	Word count	<i>must</i>		<i>should</i>		<i>ought (to)</i>	
		norm.	abs.	norm.	abs.	norm.	abs.
Austen's <i>Pride and Prejudice</i>	122,204	25.0	(305)	20.0	(245)	3.6	(44)
Brontë's <i>Wuthering Heights</i>	117,111	16.2	(190)	22.5	(263)	1.3	(15)
Dickens' <i>Great Expectations</i>	186,564	10.6	(198)	18.8	(350)	1.7	(31)
Stoker's <i>Dracula</i>	162,589	27.1	(440)	12.2	(199)	0.6	(9)
Wilde's <i>The Importance of Being Earnest</i>	20,740	17.8	(37)	23.6	(49)	0.0	(0)

Figures

You may also choose to display your data graphically in figures (graphic illustrations of the data, such as bar charts, pie charts, etc.). Figures are not a substitute for tables, but a way of illustrating even more clearly what you have already shown in a table. Never trust summary statistics alone; always visualize your data. See Figure 1, which provides information from Table 1 in a graphic form. Note that the caption is placed under each figure, whereas the caption for a table is placed above. Additionally, in figures, the coloring of variables should be consistent, e.g. red for *should* in Figures 1 and 2.





If you are not yet familiar with the technicalities of creating figures in Word or Excel, it may be a good idea to attend one of the IT courses provided by the University of Zurich (<https://www.zi.uzh.ch/en/students/ittraining.html>) or check Microsoft support materials (<https://support.microsoft.com/en-us/training>).

Normalized figures and percentages

Normalized figures (Figure 2) are necessary to compare how frequently a feature occurs in multiple data sets of different sizes. Percentages, on the other hand, are used to illustrate the frequency of a feature in relation to other features (Figure 1). When you use percentages or normalized figures, you must still give the totals on which your percentages are based. By comparing Figures 1 and 2, you can see that the use of normalized figures and percentages each tell a different story, both of which are accurate, but one might be more helpful in answering your particular research question. For example, the two figures show that, comparatively, *Great Expectations* uses the highest percentage of *should* (60%), whereas *should* is actually used more frequently per 10,000 words in *The Importance of Being Earnest* (23.6 instances per 10,000 words).

4. Statistics: Some basics

Descriptive versus inferential statistics

One reason for doing quantitative investigations of a limited text sample is often that we want to find out how something generally works in a particular variety of a language, for instance how often and when speakers of English use different types of relative pronouns, contracted forms of verbs, a particular form of negation, etc. – the range of questions is infinite. Obviously, we cannot investigate everything that has ever been said or written in English, so we have to limit ourselves to what we have reason to believe are representative samples of the language.

A representative sample is a portion of something that represents and reflects a larger set/population. If we have made such a study and think our data are representative, we can hope to infer from our results what is likely to be the case for the larger population, i.e. we can do some inferential statistics rather than just descriptive statistics. Table 1 and Figures 1 and 2 are examples of “descriptive statistics”; they illustrate quantitative differences in the form of tables and graphs. Descriptive statistics is the term given to the analysis of data that helps describe, show or summarize data in a meaningful way in order to conserve time and space. Descriptive statistics do not, however, allow us to make conclusions beyond the data we have analysed or reach conclusions regarding any hypotheses we might have made. Inferential statistics, by contrast, are used to test whether the results are powerful enough or the difference(s) are big enough that they are not due to random fluctuations. If the results are sufficiently robust, we can say that they are statistically significant – that is, true – and we can then draw some more general lessons from the study. Two commonly used significance tests in linguistics are the chi-square test and the log-likelihood test. Note that the word *significant* should only be used to describe differences in data sets when a significance test has been carried out and the results of the test are indicated.

You can learn more about statistics for linguistics by consulting the textbook and tutorials by Schneider & Lauber (<https://dlf.uzh.ch/openbooks/statisticsforlinguists/>) or Brezina (2018: <http://corpora.lancs.ac.uk/stats/>). At the MA level, the course “Methods and Theories” will provide you with solid training in statistics.

Sample size and reliability

Sample size is important. It has a direct effect on the reliability of the data. Reliability refers to how repeatable your findings are. In other words, what are the chances that you would get the same results if you were to do your study again? If you toss a coin twice and you get heads both times, you would not draw the conclusion that something was wrong with the coin. If you get heads ten times in a row, you might be suspicious that someone has tampered with the coin, and if you get heads a hundred times, you will be virtually sure that this is the case. This is common sense, and the moral is of course: the bigger your sample is, i.e. the more observations you make, the better the chances are that your sample will predict something about the language of the population you are studying (if the sample is well chosen). Because the number of observations is important, you must always tell the reader the absolute numbers in your tables, not just the proportions of different observations.

Questions arising from a limited size sample

Some interesting questions are then: How small a sample can you get away with? How can you know that differences that you have observed in your sample are not due to chance, but probably reflect real tendencies in the language? When can you know that observed differences in the use of a language form between samples consisting of, say, men and women, or speech and writing, represent real differences in language and are not just due to chance? In order to answer these questions, you need to know a little more about statistics and possibly use a few easy tests. The website <http://corpora.lancs.ac.uk/> (McEnery and Hardie 2011) provides helpful resources to complete significance tests, e.g. <http://corpora.lancs.ac.uk/stats/toolbox.php>.

If your sample is too small for reliable conclusions, one option is to increase the sample size (before you carry out your analyses). However, there are situations in which it is not possible to increase the size of a sample, for instance when studying the language of the earliest periods of English, from which only a limited number of written records have survived. In addition, statistical tests are very helpful tools if – and only if – they are understood and applied correctly. Not all statistical tests are equally applicable to linguistic data. And statistical tests can only be as reliable as the data to which they are applied. It is possible, for instance, that a test indicates a statistically significant result when, in fact, the data sample is skewed, or the categorization system is subjective or unreliable. Therefore, the most important point is that you are transparent, that you discuss limitations of your data explicitly and that you do not overstate your data, i.e. that you do not make claims that are stronger than your data allow.

5. Style sheet

Academic writing is formal and complex, yet clear, precise and objective. Regarding formal aspects, one should avoid contractions, and Latin-based verbs are preferred to phrasal verbs. Generally, if something can be expressed in fewer words, the more concise version is the more formal version.

Academic writing is often considered to be more complex than other registers. However, the goal is not to make things more complex than necessary. The topics and concepts dealt with in research are complex enough themselves, thus the goal is to describe, explain and discuss complex matters in a clear, precise and thorough manner. This often leads to the use of more pre- and post-modifiers than in other registers. Information and definitions should not be taken for granted, but rather sufficiently explained.

Maintaining an objective voice in academic writing is of the utmost importance. This does not mean that you cannot use the first-person singular pronoun. When discussing your research, specifically the steps that you took in your method, you should use the first-person singular pronoun (the first-person plural pronoun is reserved for papers completed by multiple authors). You should, however, avoid addressing your reader as *you*. If such a reference cannot be avoided, *one* may be used. Maintaining an objective and impersonal voice does not mean that your text does not contain a point of view; however, the arguments and claims you make need to be based on evidence, either from previous research or from your own study, rather than on your subjective beliefs. This evidence needs to be explicitly referenced. Such claims and arguments may also need to be hedged; in other words, the strength of the language you use should accurately represent the strength of the amount of your evidence on which your claims are based. Additionally, the goal of linguistic research is not to cast prescriptive claims, but rather to make descriptive observations.

Introducing concepts

If you introduce technical concepts in your text, you can highlight them with double quotes or capitals. Avoid italics for this purpose (italics are reserved for linguistic examples, cf. the

examples below). Quotation marks or capital letters should only be used when you introduce concepts for the first time.

When Hamlet said *words, words, words*, he used three “tokens” but only one “type”.

When Hamlet said *words, words, words*, he used three TOKENS but only one TYPE.

When introducing an abbreviation, provide the full-length description followed by the abbreviated form in parentheses. If you are introducing a database or corpus, provide its full name in italics with the abbreviation in parentheses. Always introduce abbreviations at first mention and then use the abbreviation in all subsequent cases.

Kachru’s (1985) model is often used in World Englishes (WEs) research. The data for this study are taken from the *British National Corpus* (BNC).

Emphasis

Academic prose avoids expressions of emotion and subjectivity. Do not use italics to show emphasis in your text.

RIGHT: There is not a single instance of double negation in this text.

WRONG: There is not a *single* instance of double negation in this text.

Quoting linguistic examples from your data

Letters, words, or phrases cited as linguistic examples should be italicized in the text; do not use double quotation marks for this purpose. Translations or other explanations of meaning should be given in single quotation marks:

RIGHT: The quantifier *many* means ‘a lot’.

WRONG: The quantifier “many” means (...)

Sentences quoted from your data should be set apart from the main body of the text by indentation (see (10)). This is especially important if they are longer than a few words. In this case they should be preceded by Arabic numerals in parentheses. Note that indented and numbered examples are not italicized. This is a good example of how you may proceed:

Consider the case of conversion in (10):

(10) he stopped and upped the volume (BNC AOU 1013)

All of the examples included in your text must be discussed. You need to highlight what is important, how you interpret it and indicate how these examples do or do not support your hypothesis. Notice that examples should be numbered and referred to by number, not as “the following sentence”. Additionally, you must indicate the source of your examples.

Quoting from the literature

There are three main ways of including sources in your text: direct quotes, paraphrases, and summaries.

(1) Direct quotes. When you quote a text directly, you repeat the exact wording as used in the original and mark word or text by using “double-quotation marks”, thus:

As Tannen (1991: 143) states, “the dissatisfaction may be mutual” if men do all the talking and women always have to listen.

Quotations longer than three lines should be indented and set apart from the text. Notice also that indented quotations should not be surrounded by double quotes, as in the following example. Additionally, if you need to alter the quote to make it grammatically fit into your sentence/text, the altered text needs to be enclosed in square brackets []. If you remove part of the quote, this should be indicated by square brackets and ellipsis [...].

In her study of misunderstandings between men and women in conversation, Tannen (1991: 143) states:

If women are dissatisfied with always being in the listening position, the dissatisfaction may be mutual. That a woman feels she has been assigned the role of silently listening audience does not mean that a man feels he has consigned her to that role—or that he necessarily likes the rigid alignment either.

or

In her study of misunderstandings between men and women in conversation, Tannen states that

[...] the dissatisfaction may be mutual. That a woman feels she has been assigned the role of silently listening audience does not mean that a man feels he has consigned her to that role—or that he necessarily likes the rigid alignment either (1991: 143).

(2) Paraphrase. In many cases, it is more suitable to paraphrase a text rather than to use a direct quote; this means that you reproduce the meaning of the original in your own words. Note that it is not enough to simply change the word order or to replace individual words by synonyms. Anything that is too close to the original may still constitute plagiarism. Do not forget to add a reference to paraphrased content. For example:

... Tannen (1991: 143) states that if men do all the talking and women always have to listen, both genders could end up dissatisfied. ...

(3) Summary. Finally, you may choose to summarize entire articles, chapters, or even books. This is often done when pointing out research gaps, e.g. when you want to show what has already been done without going into a lot of detail. In such cases, it is usually enough to use basic forms of referencing:

... numerous studies on the social and historical importance of this community (Hsu 1971; Lyman 1974; Nee and Nee 1986; Takaki 1989; Tsai, Ying and Lee 2000; Wong 2006) as well as on Chinese-American literature (Wu 1983; He 1996; Wong 1997; Yin 2000) have been published.

Note that semi-colons (;) are used between each different author and that the sources are ordered chronologically. If you cite multiple studies from one author, these are separated by comma (e.g. Jucker 2009, 2014).

References in the text

Whenever you discuss information that you have found in other sources, you need to acknowledge the source and add a citation.

Use the following format for references in your text. Do not use footnotes for references, and do not use *ibid.* for repetitions of the same source.

Chomsky (1980a: 3) introduced the term ...

Many syntacticians (e.g. Matthews 1981: ch 3) distinguish ...

For sources written by more than three authors you can use *et al.*, as illustrated below (provide all names in the reference section):

To investigate the effect of ethnic orientation on the use of linguistic features, Nagy et al. (2014: 11–13) ...

When you cite two or more sources, list them chronologically and separate them with a semi-colon, as in the following example:

This insight has gained ground in recent years (Atlas and Levinson 1981; Sperber and Wilson 1985).

You may sometimes want to quote a source which is known to you only through another publication because the original is not available. You can do it like this:

A collocation can be defined as “actual words in habitual company” (Firth 1957: 14, quoted in Kennedy 1998: 108).

Both references must appear in the reference section at the end of your paper. Use this form of quoting sparingly and always make sure that you really do not have access to the original source.

When introducing a definition from a dictionary, e.g. the *Oxford English Dictionary*, use the following format:

Originally, the definition of diaspora is related to the expulsion of the Jews from Babylonia and refers to

[t]he body of Jews living outside the land of Israel; the countries and places inhabited by these, regarded collectively; the dispersion of the Jewish people beyond the land of Israel (OED, s.v. *diaspora*, n.).

In this example, “s.v.” stands for abbreviated Latin *sub verbo* ‘under the word’ and refers to the entry on the noun (n.) *diaspora* in the OED.

Formatting the reference section

Include only works cited. Do not include works that you may have consulted or even read in full, but which are not cited in the paper.

The year of publication is typed immediately after the name(s) of the author(s). This makes it easier for the reader to check your references in the text. Titles of books and journals are

italicized, and upper case is used (headline style). Titles of articles and chapters are not italicized but have quotation marks around them and upper case is used for content words. Apply your formatting consistently, even if it differs from the original formatting of your source(s). For books (but not for journals), indicate place of publication and publisher at the end. Remove all series information for monographs from the references. For place information, only include the city, e.g. “Malden, M.A.”, also remove double locations e.g. “Amsterdam and New York: John Benjamins”. Give page references for articles and sections of books consisting of collections of articles. Include number and issue information for journals that appear more than one a year, e.g. *Lingua* 60.1: 30–50 or *Lingua* 60(1): 30–50. As far as possible, give full first names of authors, as they facilitate bibliographical searches.

Sources/Primary sources

Include data sources you used that consist of published material (e.g. newspapers, dictionaries, films, electronic corpora, websites). Always list the title of the publication in italics (with its abbreviation in parentheses, if available), provide the URL if the source is available online, and add the date of access. If available/applicable, also list the author/compiler, the publisher, and the date of publication. For corpora, you may find useful information at <https://varieng.helsinki.fi/CoRD/index.html>.

If you created your own data set, sampling for instance episodes or news articles, you should provide the general source of your data (e.g. the TV series if you used several episodes or the newspaper if you sampled several articles). Your sampling method should be detailed in the data section outlined above.

Corpora:

The Corpus of Contemporary American English (COCA). Compiled by Mark Davies. 2008–. Available at <<http://corpus.byu.edu/coca/>> (accessed September 19, 2018).

British National Corpus (BNC). 1994. Available at <<http://www.natcorp.ox.ac.uk>> (accessed September 19, 2018).

TV/movie transcripts:

How I Met Your Mother. Transcripts available at <https://www.springfieldspringfield.co.uk/episode_scripts.php?tv-show=how-i-met-your-mother> (accessed September 19, 2018).

Books/speeches with an author:

Wilde, Oscar. 1895. *The Importance of Being Earnest*. Available at <<https://www.gutenberg.org/ebooks/844>> (accessed September 19, 2018).

Obama, Barack H. 2010. *First Presidential State of the Union Address*. Available at <<https://www.americanrhetoric.com/speeches/stateoftheunion2010.htm>> (accessed September 19, 2018).

Websites without authors:

"What video game has the best open world?". 2017. *Reddit*. Available at <https://www.reddit.com/r/AskReddit/comments/6t8got/what_video_game_has_the_best_open_world/> (accessed September 19, 2018).

The New York Times. 23 March – 29 March 2017. Available at <<https://www.nytimes.com>> (accessed March 29, 2017).

Print references/Secondary sources

Books

Use the following format; notice the indentation of the second lines of references.

Single author:

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.

Multiple authors:

Jucker, Andreas H., and Irma Taavitsainen. 2013. *English Historical Pragmatics*. Edinburgh: Edinburgh University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. A *Comprehensive Grammar of the English Language*. London: Longman.

Edited volumes:

Wanner, Eric, and Leila R. Gleitman, eds. 1982. *Language Acquisition: The State of the Art*. Cambridge: Cambridge University Press.

For the second or revised edition of a book you should give the date of the new edition:

Romaine, Suzanne. 1995. *Bilingualism*. (2nd ed.). Malden: Blackwell.

Gleason, Henry A. Jr. 1961. *An Introduction to Descriptive Linguistics*. Revised edition. New York: Holt, Rinehart and Winston.

Provide volume information after the title:

Booij, Geert E., Christian Lehmann, and Joachim Mugdan, eds. 2000. *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*. Vol. 1. Berlin: De Gruyter.

Wells, John C. 1982. *Accents of English*. Vol. 2: *The British Isles*. Cambridge: Cambridge University Press.

When quoting reprints of older publications, you should indicate both the original and new publication date:

Bühler, Karl. [1934] 1978. *Sprachtheorie: Die Darstellungsfunction der Sprache*. Frankfurt: Ullstein.

Articles in journals and collected volumes

Article in a journal:

For journal articles, provide volume number.

Deterding, David. 2005. "Emerging Patterns in the Vowels of Singapore English". *English World-Wide* 26: 179–197.

Article printed in a collection:

Labov, William. 1972. "Rules for Ritual Insults". In David Sudnow, ed. *Studies in Social Interaction*. New York: The Free Press, 120–169.

If the collective volume is also entered in your list of references under the name of the editor, which is a good idea if you quote several articles published in the same volume, you can use the following format:

Labov, William. 1972. "Rules for Ritual Insults". In David Sudnow, ed., 120–169.

Sudnow, David, ed. 1972. *Studies in Social Interaction*. New York: The Free Press.

In this case, you only refer to the collective volume once in the relevant section of your alphabetical list of references.

Page references in the text are made to the edition you used, for example: Nickel (1978: 150). Use one of the following formats for reprinted articles:

Nickel, Gerhard. [1968] 1978. "Kontextuelle Beziehung zwischen Sätzen im Englischen". In Wolfgang Dressler, ed. *Textlinguistik*. Darmstadt: Wissenschaftliche Buchgesellschaft, 147–166.

Nickel, Gerhard. 1968. "Kontextuelle Beziehung zwischen Sätzen im Englischen". Praxis des neusprachlichen Unterrichts 15: 15–25. Also in: Wolfgang Dressler, ed. 1978. *Textlinguistik*. Darmstadt: Wissenschaftliche Buchgesellschaft, 147–166.

Two or more publications by the same author in the same year should be distinguished by adding a, b, c, etc. (also for the in-text citation):

Kohnen, Thomas. 2008a. "Directives in Old English: Beyond Politeness?" In Andreas H. Jucker and Irma Taavitsainen, eds. *Speech Acts in the History of English*. Amsterdam: Benjamins, 27–44.

Kohnen, Thomas. 2008b. "Linguistic Politeness in Anglo-Saxon England? A Study of Old English Address Terms". *Journal of Historical Pragmatics* 9: 140–158.

Articles in proceedings

Lewandowski, Natalie, Matthias Jilka, Guiseppina Rota, Susanne Reiterer, and Grzegorz Dogil. 2007. "Phonetic Convergence as a Paradigm of Showing Phonetic Talent in Foreign Language Acquisition". *Cognitive Science 2007: Proceedings of the 8th Annual Conference of the Cognitive Science Society of Germany*, Saarbrücken, 1–7.

Dictionaries and encyclopedias

In the case of dictionaries, the editor's name can be omitted or, if given, should follow the title of the dictionary, not precede it:

The American Heritage Dictionary. 1969. William Morris, ed. Boston: Houghton-Mifflin.

A similar approach can be applied to encyclopedias:

The Wiley Blackwell Encyclopedia of Anglo-Saxon England. 2014. Michael Lapidge, John Blair, Simon Keynes, Donald Scragg, eds. Malden: Wiley Blackwell.

If the encyclopedia entries are anonymous, you should format as above. If, however, they are signed, refer to the individual entries explicitly. With multiple entries from the same encyclopedia, you can use the same shortened format as in collections of articles:

Keynes, Simon. 2014. “Ælfric”. In Michael Lapidge et al., eds., 9.

When you use several lexical entries from the same dictionary, it is enough to refer to the dictionary only once in your bibliography. In the main text, use short references as explained above, e.g. (OED, s.v. diaspora, n.).

Unpublished work

Below is an example of a reference to an unpublished doctoral thesis (note: no italics). Avoid referencing unpublished student work below PhD level (e.g. BA or MA theses).

Johnson, Ellen. 1992. “Lexical Change and Variation in the Southeastern United States in the Twentieth Century”. Ph.D. Dissertation, University of Georgia.

For unpublished conference papers, use the following reference format (note: no quotation marks).

Beal, Joan, and Lourdes Burbano-Elizondo. 2010. ‘All the Lads and Lasses’: Lexical Variation in Tyne and Wear. Paper presented at Sociolinguistics Symposium 18, University of Southampton.

Online references

Online books

Hauben, Michael, and Ronda Hauben. 1995. Netizens: On the History and Impact of Usenet and the Internet. <<http://www.columbia.edu/~hauben/netbook/>> (accessed January 10, 2012).

Online articles in journals and newspapers

Article in an online journal:

This citation format only applies to journals that are exclusively available online; if a print publication is available, please use the citation format for articles in journals above.

Waldvogel, Joan. 2007. “Greetings and Closings in Workplace Email”. *Journal of Computer-Mediated Communication* 12 <<http://jcmc.indiana.edu/vol12/issue2/waldvogel.html>> (accessed January 10, 2012).

Article in a newspaper:

If information on the author or the exact date of publication is not available, use n.a. and n.d., respectively.

Broder, John M., and Michael Powell. 2008. "McCain Defeats Romney in Florida Vote". The New York Times. 30 January 2008 <<http://www.nytimes.com/2008/01/30/us/politics/30florida.html?r=1&hp&oref=slogin>> (accessed March 12, 2014).

Websites

Population Statistics Stocktake: 2011. 2012.

<www.stats.govt.nz/browse_for_stats/population/pop-stats-stocktake.aspx>
(accessed February 24, 2017).

Curzan, Anne. 2013. Slash: Not just a Punctuation Mark Anymore. The Chronicle of Higher Education, (Lingua Franca). <www.chronicle.com...> (accessed November 10, 2016).

Include the reference to an author if that information is available.

Online dictionaries

Merriam-Webster Online. <<http://www.merriam-webster.com/>> (accessed December 1, 2011).

Oxford English Dictionary Online. 2023. Oxford University Press. <<http://www.oed.com>> (accessed July 13, 2017).

For additional online sources, databases, and CD-ROMs, see also Kavka/Schweizer et al. (2022) or consult the stylesheet available at <https://benjamins.com/series/eww/stylesheet.pdf>.

6. Checklist for papers

When you have finished writing your paper, read through it critically multiple times. In one round, focus on content and, in another, on language. Then ask one or two fellow students to read through your paper before you hand it in. This will benefit you as well as the person reading the paper. The more familiar you are with a piece of writing the more difficult it is to see your own mistakes, redundancies, or gaps in information.

Before handing in your paper, check your paper for aspects included in Section 6 as well as the following:

- Consistent use of either American or British English in spelling and grammar;
- Appropriate and consistent use of tenses;
- No pagination for Table of contents;
- Consistent referencing style, citation of all references, and inclusion of all and only the references cited in your text;
- Consistent use of font style in text, headings, tables, and captions.

References

- A.P.A. 2017. "Ethical Principles of Psychologists and Code of Conduct". <http://www.apa.org/ethics/code/index.aspx> (accessed August 31, 2023).
- BAAL. 2021. "Recommendations on Good Practice in Applied Linguistics". <https://www.baal.org.uk/wp-content/uploads/2021/03/BAAL-Good-Practice-Guidelines-2021.pdf> (accessed August 31, 2023).
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Brown, Keith, and Jim Miller. 2013. *The Cambridge Dictionary of Linguistics*. Cambridge: Cambridge University Press.
- Bussmann, Hadumod. 1996. *Routledge Dictionary of Language and Linguistics*. New York: Routledge.
- Butler, Christopher. 1985. *Statistics in Linguistics*. Oxford: Blackwell.
- Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell. (6th edition)
- D'Arcy, Alexandra, and Taylor Marie Young. 2012. "Ethics and Social Media: Implications for Sociolinguistics in the Networked Public". *Journal of Sociolinguistics* 16: 532–546.
- Jackson, Howard. 2007. *Key Terms in Linguistics*. Continuum, London.
- Kavka, Mischa, and Bernard Schweizer. 2006. *Literature Guide: Essay Writing, Presentations, Research*. 4th Revised Edition by Morgane Ghilardi, Alan Mattli, and Hannah Schoch. 2022. Zurich: English Department, University of Zurich. <https://www.es.uzh.ch/en/studies/studyadvice/furtherinformation/downloads.html> (accessed August 31, 2023).
- Labov, William. 1972. *Sociolinguistic Patterns*. Oxford: Blackwell.
- Matthew, Peter H. 2014. *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- Schneider, Gerold and Max Lauber. *Statistics for Linguistics: A patient, slow-paced introduction to statistics and to the programming language R. An interactive E-book with 50 movies and extensive exercises*. University of Zurich. <https://dlf.uzh.ch/openbooks/statisticsforlinguists/> (accessed August 31, 2023).
- Schreier, Daniel. 2013. "Collecting Ethnographic and Sociolinguistic Data". In Manfred Krug, and Julia Schlüter, eds. *Research Methods in Language Variation and Change*. Cambridge: Cambridge University Press, 17–35.
- Wray, Alison, and Aileen Bloomer. 2006. "Plagiarism and How to Avoid It". In Alison Wray, and Aileen Bloomer, eds. *Projects in Linguistics. A Practical Guide to Researching Language*. London: Hodder Education, 236–251.