

Writing a linguistics paper¹
George Aaron Broadwell
April 2008

1 General organization and content

1.1 Organization

Your paper should say clearly at the beginning what the goal of the paper is, and should be organized into sections that follow some logical sequence. At the end of your paper, you should write a conclusion.

When writing a linguistics paper avoid “travelogue” style in which you narrate the order in which you discovered things, or the different theories you had about a phenomenon. Just give us the facts – we’re not so concerned with how you discovered them, what was difficult or confusing, etc.

1.2 Examples

A general principle for papers in linguistics is that every claim about the language should be exemplified. If you say, e.g. that “Body part nouns must have the /x-/ prefix” then you must give examples that show that. If you say that something is impossible or ungrammatical, you need to cite examples that show that.

Examples may also be needed if your statement implies something about the grammaticality or ungrammaticality of certain kinds of sentences. If you say something like “Objects must appear before the verb”, then this statement implies that sentences with objects after the verb are ungrammatical, so you should give an example that shows this.

When giving ungrammatical examples, you need to be sure that the only reason for their ungrammaticality is a violation of the principle/condition/rule that you are discussing. You must be careful that everything else in the sentence is correct and obeys all the other rules of the grammar. If there is some other plausible explanation for the ungrammaticality of the example you cite, then you will not persuade a skeptical reader (or your professor).

For example, suppose you want to claim that a dative subject cannot be questioned, and cite an ungrammatical sentence that shows this. If your example also contains a subject agreement violation, or has the wrong word order, or uses a verb with the wrong subcategorization, then the ungrammaticality might be due to that factor, rather than to a violation of the restriction on questioning dative subjects.

1.3 Diagrams

Many papers in linguistics will need some sort of visual aid in order to make the argument clear to the reader. I’ll call these diagrams (tree structures, attribute-value matrices, optimality-theoretic tableaux, spectrograms, etc.). Deciding how many diagrams are needed and where they

¹ Many thanks to Lee Bickmore and Pamela Munro for their comments on various versions of these guidelines. Sections 2 and 3 are adapted from Pamela Munro’s *Some guidelines regarding linguistic style*.

should go is a tricky question, and writers may disagree on some cases. However, most writers would agree that not every example in the text needs an accompanying diagram – that is overkill. A more frequent problem is not enough diagrams. Try to imagine talking about your ideas with a person while standing in front of a blackboard or overhead projector. How many diagrams would you need to put up in order for your listener to understand you?

When you include diagrams, you should number them, and the accompanying text should refer to the diagram number. It should be clear how a cited example corresponds to a given diagram. For example, if you use a tree diagram, then the words in the tree must exactly match up with those in some cited example.

If your word-processor allows you to add captions, these are often helpful in letting the reader understand what the diagram shows and how it connects to the argument made in the paper.

1.4 Integrity of citation

Don't present other people's ideas as your own. If some idea is taken from a written work, you must say so and cite the work. If some idea or bit of information came from another person (including one of your professors), you should cite them as well. The usual way is as follows:

Ken Hale (p.c.) suggests that the Winnebago pronoun *nee* has this interpretation.

There are sometimes gray areas as well. Citations for personal communication usually reflect isolable, specific issues of data or theory. It is less common to cite someone for pointing out another way of thinking about the problem or a possible alternate interpretation, though it would not be incorrect to do so.

See below for the stylistic details of how citations in linguistics papers should be formatted.

1.5 Integrity of data

Don't make up data. If the data you need to demonstrate some point are unavailable, you have to honestly admit this in the paper.

If you use data from other people's works, you need to cite the source. It is acceptable to make very minor adjustments to other people's data, so long as you say you are doing so. The general practice seems to be that it is acceptable to modify the orthography and the glossing to match what is used elsewhere in your paper. But you cannot change the words of a cited example.

If the example you want to cite is very long, it is possible to cite only the portion that is relevant, but in doing this you must be scrupulously fair and not omit any part of the cited example that is important evidence either for or against the claims you make in the paper. For example, if you claimed that experiencers always receive dative case, it would be unethical to cite a long example from another source, omitting the crucial examples of experiencers with another case.

If you have data in your own notes that bears on a proposal you are making, you need to acknowledge it. For example, if you claimed that experiencers always receive dative case, but you have two counterexamples in your notes, then it is unethical to write the paper, while failing to mention the problematic cases. An honest linguist will double-check the facts and if they are truly counterexamples, then either revise the theory, come up with an explanation of the counterexamples, or acknowledge the problem in a footnote.

Similarly, if other authors present data that represent a counterexample to your claim, you

need to cite and discuss these data. An honest linguist doesn't just ignore data that challenge his or her theory.

2 Format of papers

A formal paper in linguistics needs to contain several elements in addition to the text itself.

2.1 Acknowledgments

In a footnote to the title or to the first sentence, you conventionally thank any people who have helped you in writing it. This especially includes speakers of other languages who have served as consultants for the paper. I'd advise against acknowledging more spiritual/emotional kinds of assistance, so it's generally inappropriate to thank your parents, your friends, the Buddha, etc. (About the only place this rule is not followed is in dissertations, where your acknowledgments section can be as extensive and maudlin as you want -- "And to my teddy bear Humphrey, who patiently endured my seven years of research on the uvula ...")

You also don't thank people who have influenced your thought but who you haven't met (e.g. Noam Chomsky, Jacques Derrida, Franz Boas).

2.2 Orthography

If you're writing about an unfamiliar language, you need to explain the orthography you're using. Typically, you will explain the mapping of orthographic symbols (contained in angled brackets) to phonetic or phonemic symbols:

- 1) In the orthography used here, <ch> represents [tʃ] and <sh> is [ʃ].

There is nothing wrong with using practical orthographies or digraphs in your paper, so long as a footnote explains them.

In languages written in a non-Roman orthography, you need to explain your transcription system:

- 2) In the transliteration of Greek used here, ω is [ō].

2.3 **List of abbreviations** This can also be included in the first footnote. Abbreviations are used extensively in interlinear glossing (see below).

2.4 References

Style varies somewhat, but I'd suggest the following format for bibliographic references:

Goddard, Ives. 1974. Remarks on the Algonquian independent indicative. *International Journal of American Linguistics* 40:317-27.

- , 1979. *Delaware Verb Morphology: A Descriptive and Comparative Study*. New York: Garland.
- , 1987. Fox participles. *Native American Languages and Grammatical Typology*, ed. Paul D. Kroeber and Robert E. Moore, pp. 105-118. Bloomington, IN: Indiana University Linguistics Club.

The first example shows an article in a journal, the second a book, and the third an article in an edited volume. Note that in the titles of articles, only the first word and proper names are capitalized. In the names of books and journals, every significant word is capitalized.

It is important that the citations in the paper and the items in the bibliography match exactly. Do not include items in the bibliography that are not cited in the paper, and do not cite references which are not listed in the bibliography.

2.5 Optional elements

Larger manuscripts may also contain tables of content, lists of figures, and appendices.

3. Style

Papers and articles in linguistics typically adopt stylistic features which are not found in other types of writing, partially because we need to distinguish the language that is the subject of the paper from the surrounding text. Here are some of the most commonly accepted stylistic conventions:

1. **Citing other authors** The standard way to cite previous authors is as follows:

3 Chomsky (1981) argued that ...

Reference to a particular page or pages is indicated by a colon

4 Chomsky (1986:38) claimed that ...

When an author publishes more than one item in a given year, the items are distinguished by small roman letters:

5 Goddard (1987a) proposed that ...

When the genitive is used with such a reference, the /'s/ precedes the year:

6 Chomsky's (1981) definition of governing category ...

When citations to previous literature occur inside parentheses, the parentheses around the year are often omitted:

7 Chomsky (1982) claimed that ... (following suggestions in Rizzi 1980).

Let me stress once more that it is absolutely essential to acknowledge ideas or text that you've taken from other works. Failure to do so is plagiarism.

2. **Citing linguistic material**

A. **Underlining in the text** Cited forms are set off from the surrounding text whenever they are used in a sentence. This is often done by underlining:

8 The Choctaw word waak is a borrowing from Spanish.

9 The past tense of bite is bit.

As the second example shows, this is done even when the cited forms are English.

If italics are available, they may also be used for this purpose:

10 The Zapotec word *bini* ends in a voiceless vowel.

Don't alternate underlining and italics in your paper -- choose one and stick to it.

You also use italics or underlining with the titles of books and journals:

11 In *Aspects*, Chomsky argued against treating grammatical relations as primitives.

In your final bibliography, some styles use italics or underlining for the same purpose. (Though journals vary considerably in their rules here.)

B. **Single quotes in the text** Glosses (i.e. translations) for data are given in single quotation marks. Unless sentence structure warrants otherwise, glosses follow the cited data immediately, with no other punctuation:

12 Choctaw *oklab homma* 'red people' was the source for the name of the state *Oklahoma*.

13 The Chickasaw cognate to Choctaw *oklab*, however, means 'town', not 'people'.

Note that, as the last example shows, it is usual for punctuation to follow the single quotation mark of a gloss. (Outside of linguistic style, recall that punctuation generally precedes a quotation mark.)

Single quotes are used for glosses even if the cited word does not appear:

14 The verb 'be' is irregular in many languages.

There is no reason to enclose the names of grammatical morphemes in quotes unless their gloss is actually their meaning:

- 15 ... the Choctaw complementizer *-kat...*
- 16 ... the Pima auxiliary elements *-ñ* (first person singular), *-p* (second person singular), ...
- 17 ... Chickasaw *-hookeya* 'but'...

Normally either a preceding descriptive phrase or a following one in parentheses is used to identify grammatical morphemes.

C. **Phonemic slashes in the text** Another option for citing language material in the text is the use of phonemic slashes:

- 18 /-akili/ is a frequent emphatic particle in Choctaw.

This is sometimes a useful option when italics are unavailable and underlining is being used for some other purposes in the orthography (e.g. to mark vowel nasalization, low tone, laryngealization, etc.).

When using slashes, however, it is important that the material between slashes be in a phonemic orthography. For European languages with Roman orthographies, it will often be more appropriate to cite linguistic material in the original orthography unless some point about the phonology is being made:

- 19 In Spanish the indefinite article is omitted following the phrase *a manera de*.
- 20 In Castillian Spanish *acerca* is pronounced /aθerka/.

In languages in which the orthography is phonemic, one may use phonemic slashes in both phonological and non-phonological contexts. However, this device is best used sparingly, to set off single words or morphemes. It is tedious when used for sentences:

- 21 The Chatino sentence /ndya'a ndijni' cñii cha' cui'ya cube'/ 'he went to ask to borrow money to buy a boar' shows a purpose clause preceded by /cha'/.

A better way to cite this sentence is to use interlinear glossing:

- 22 The following Chatino sentence shows a purpose clause preceded by /cha'/:

1.)	Ndya'a	ndijni'	cñii	cha'	cu-i'ya	cuba'.
	com:go	borrow	money	in:order:to	pot-buy	boar

'He went to borrow money in order to buy a boar.'

See the section on interlinear glossing below for more details on this way of citing foreign language data.

Authors differ on whether it is acceptable to use non-IPA (International Phonetic Alphabet) symbols

between phonemic slashes. In the example just given, the orthographic sequence <ch> corresponds to [tʃ]. I think that non-IPA phonemic symbols are acceptable, so long as it is clear to your reader how the symbols between the slashes are to be interpreted.

D. Interlinear glossing

The best way to cite data of more than a few words is to use interlinear glossing. Interlinear glossing involves the use of three lines of text -- the original broken up into morphemes, a morpheme-by-morpheme gloss of the original, and a free translation, as in the following example from Ge'ez:

23	Wa- <u>h</u> aywa	zentu	'ab	za-tarfa	'em-mawā`eli-hu
	and-lived	this	father	rel-passed	from-days-3s
	ba-hed'at	wa-salām.			
	in-calmness	and-peace			

`And this father lived and died (lit. passed from his days) in calmness and peace.'

The following points are important: (i) the number of hyphens and spaces in the original must match the number of hyphens and spaces in the gloss line; (ii) the beginning of every word in the original must line up with the beginning of every word in the gloss line. Try to make sure that all the parts of an example with interlinear glossing stay on the same page; it is hard to read the examples otherwise.

This means that when a single word in the original corresponds to more than one word in the gloss line, some special device must be used. I prefer the use of colons between the words of the gloss, as in the gloss for the word *cha'* in the following:

24	Ndya' <u>a</u> ndijni'	cñii	cha'	cu-i'ya	cuba'.
	com:go borrow	money	in:order:to	pot-buy	boar

`He went to borrow money in order to buy a boar.'

Colons can also be used when it is not necessary or possible to break a word up into its constituent morphemes. This is the case in the gloss for the word *ndya'a*, which is an irregular completive of the verb `to go' (comparable to our *went*).

Alignment of words can be a problem when you are working with certain kinds of fonts. On many computers, things that look like they are aligned on screen will print out differently because most fonts are proportionally spaced (i.e. different letters take up different amounts of space on the page). There are two solutions to the alignment problem: a) use a non-proportional font, such as Courier, or b.) make every word of the original correspond to a tab, and use the tabs to align the material in the gloss line. You can do whichever of these seems easier. It is acceptable to shift fonts between the text and the examples, so a paper with text in New York and examples in Courier is okay.

Gloss lines typically include lots of abbreviations which are identified elsewhere in the paper. Note that sentences from other languages generally begin with a capital letter and end with a

period (or other punctuation as appropriate). This principle may be overridden if capital letters serve some other purpose in the orthography being used. When citing words or phrases, there is no need to capitalize.

Always provide a translation of every cited example. It is silly and annoying to leave examples untranslated because 'everyone knows French (or Greek, or Menomini)'. If the original involves some idiomatic material, it is often wise to include some acknowledgment of this, as in the previous example where *pass from one's days* is a Ge'ez idiom for 'die'.

3. Other kinds of citation

A. Capitals in the text In discussions of semantics, putative semantic translations or semantic universals are sometimes written in capitals:

25 I will assume that the semantic decomposition of the verb redden is CAUSE (x, BECOME (y, red))...

B. Double quotes in the text Double quotes may be used to set off words and phrases whose use has some particular theoretical sense (possible associated with a given school or scholar) which the author wants to insist on. Such uses of double quotes are best if the author makes it clear whose sense of the term is being used:

26 I will use the term "diathesis alternation" (Hale and Keyser 1986) to describe alternations in predicate valence.

27 "Binding", in the sense of Chomsky (1981), refers to relations between antecedents and anaphors.

Once again, punctuation follows quotation marks of this sort. Once this kind of terminological preliminary is completed, it is often helpful to discontinue the use of such double quotes through the rest of the paper, since they can be distracting or seem pretentious. There is no need to use double quotes around standard linguistic terminology, unless a new definition is offered for it in the paper (and not always then). There is never any need to use double quotes around slang or cute expressions. If they are worth using at all, they are worth using without quotes.

Another use of double quotes is for citation of brief quotations from another author.

28 Nicklas writes, "You will learn to lengthen them automatically" (*Reference grammar of the Choctaw language*, p. 3), but he does not say how.

A final use of double quotes is for the citation of the titles of articles:

29 Chomsky's "Remarks on nominalization" was influential in turning the tide against generative semantics.

I'd advise against using this style of citation, since it is less clear to the reader how to follow this reference in the bibliography. A better way is as follows:

Chomsky (1970) was influential in turning the tide against generative semantics.