

Leipzig Glossing for Typst

Introduction

Interlinear morpheme-by-morpheme glosses are common in linguistic texts to give information about the meanings of individual words and morphemes in the language being studied. A set of conventions for these glosses, called the “Leipzig Glossing Rules”, was developed to give linguists a general set of standards and principles for how to format these glosses. The most recent version of these rules can be found in pdf form at [this link](#), provided by the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology.

There is a staggering variety of LaTeX packages designed to properly align and format glosses (including gb4e, ling-macros, linguex, expex, and probably even more). These modules vary in the complexity of their syntax and the amount of control they give to the user of various aspects of formatting. The typst-leipzig-glossing module is designed to provide utilities for creating aligned Leipzig-style glosses in Typst, while keeping the syntax as intuitive as possible and allowing users as much control over how their glosses look as is feasible.

This pdf will show examples of the module’s functionality and detail relevant parameters. For more information or to inform devs of a bug or other issue, visit the module’s github repository [neunenak/typst-leipzig-glossing](#).

Contents

1 Basic glossing functionality	1
2 Styling lines of a gloss	3
3 Spacing between items	5
4 Line spacing	5
5 Gloss indentation	5
6 Item alignment	6
7 Nlevel glossing: alternate syntax	7
8 Per-item Styling	8

1 Basic glossing functionality

As a first example, here is a gloss of a text in Georgian, along with the Typst code used to generate it:

from “Georgian and the Unaccusative Hypothesis”, Harris, 1982

ბავშვი-ი ატირდა
bavšv-i aṭirda
child-NOM 3S/cry/INCHO/II

“The child burst out crying”

```
#gloss(  
  header_text: [from "Georgian and the Unaccusative Hypothesis", Harris, 1982],  
  source_text: ([ბავშვი-ი], [ატირდა]),  
  source_text_style: (item) => text(fill: red)[#item],  
  transliteration: ([bavšv-i], [aṭirda]),  
  morphemes: ([child-#smallcaps[nom]], [3S/cry/#smallcaps[incho]/II]),  
  translation: ["The child burst out crying"],  
)
```

The `#gloss` function has three pre-defined parameters for glossing levels: `source_text`, `transliteration`, and `morphemes`. It also has two parameters for unaligned text: `header_text` for text that precedes the gloss, and `translation` for text that follows the gloss.

If one wishes to add more than three glossing lines, there is an additional parameter `gloss_lines` that can take a list of arbitrarily many more glossing lines, which will appear below those specified in the aforementioned parameters:

Hunzib (van den Berg 1995:46)

<i>oʒɖuʒ</i>	<i>xoʰxe</i>	<i>mykʰep</i>
oʒdig	χõχe	muq'er
oʒ-di-g	xõxe	m-uq'e-r
boy-OBL-AD	tree(G4)	G4-bend-PRET
at boy	tree	bent

“Because of the boy, the tree bent.”

```
#gloss(
  header_text: [Hunzib (van den Berg 1995:46)],
  source_text: ([oʒdig], [xo#super[n]xe], [mykʰep]),
  transliteration: ([oʒdig], [χõχe], [muq'er]),
  morphemes: ([oʒ-di-g], [xõxe], [m-uq'e-r]),
  gloss_lines: (
    ([boy-#smallcaps[obl]-#smallcaps[ad]], [tree(#smallcaps[g4]]),
    [#smallcaps[g4]-bend-#smallcaps[pret]]),
    ([at boy], [tree], [bent]),
  ),
  translation: ["Because of the boy, the tree bent."]
)
```

To number gloss examples, use `#numbered_gloss` in place of `gloss`. All other parameters remain the same.

(1) Indonesian (Sneddon 1996:237)

<i>Mereka</i>	<i>di</i>	<i>Jakarta</i>	<i>sekarang.</i>
they	in	Jakarta	now

They are in Jakarta now

```
#numbered_gloss(
  header_text: [Indonesian (Sneddon 1996:237)],
  source_text: ([Mereka], [di], [Jakarta], [sekarang.]),
  morphemes: ([they], [in], [Jakarta], [now]),
  translation: "They are in Jakarta now",
)
```

The displayed number for numbered glosses is iterated for each numbered gloss that appears throughout the document. Unnumbered glosses do not increment the counter for the numbered glosses:

Kinyarwanda

mú-kòrà
SBJV\1PL-work

“that we work”

(2) West Greenlandic (Fortescue 1984:127)

palasi=lu niuirtur=lu
priest=and shopkeeper=and

“both the priest and the shopkeeper”

2 Styling lines of a gloss

Each of the aforementioned text parameters has a corresponding style parameter, formed by adding `_style` to its name: `header_text_style`, `source_text_style`, `transliteration_style`, `morphemes_style`, and `translation_style`. These parameters allow you to specify formatting that should be applied to each entire line of the gloss. This is particularly useful for the aligned gloss itself, since otherwise one would have to modify each content item in the list individually.

In addition to these parameters, Typst’s usual content formatting can be applied to or within any given content block in the gloss. Formatting applied in this way will override any contradictory line-level formatting.

This text is about eating your head.

I’m *eat-ing* *your* *head*
1SG.SBJ=to.be eat-PROG 2SG.POSS head

I’m eating your head!

```
#gloss(  
  header_text: [This text is about eating your head.],  
  header_text_style: text.with(weight: "bold", fill: green),  
  source_text: (text(fill:black)[I'm], [eat-ing], [your], [head]),  
  source_text_style: text.with(style: "italic", fill: red),  
  morphemes: ([1#sg.#sbj\=to.be], text(fill:black)[eat-#prog], [2#sg.#poss],  
  [head]),  
  morphemes_style: text.with(fill: blue),  
  translation: text(weight: "bold")[I'm eating your head!],  
)
```

To apply styles to lines in the `gloss_lines` parameter, there is a `line_styles` parameter. This parameter takes a list of styles and provides them in order to the lines provided by `gloss_lines`. If fewer styles are provided than there are lines in `gloss_lines`, the remaining lines will be displayed with whatever the default formatting is.

(3) Lezgian (Haspelmath 1993:207)

<i>Gila</i>	<i>abur-u-n</i>	<i>ferma</i>	<i>hamišaluğ</i>	<i>güğüna</i>	<i>amuq'-da-č.</i>
now	they-OBL-GEN	farm	forever	behind	stay-FUT-NEG
now	their	farm	forever	behind	will not stay
τώρα	τους	αγρόκτημά	για πάντα	πίσω	δεν θα παραμείνει

Now their farm will not stay behind forever.

```
#numbered_gloss(  
  header_text: [Lezgian (Haspelmath 1993:207)],  
  source_text: ([Gila], [abur-u-n], [ferma], [hamišaluğ], [güğüna], [amuq'-da-  
č.]),  
  morphemes: ([now], [they-#obl\-#gen], [farm], [forever], [behind], [stay-  
#fut\-#neg]),  
  gloss_lines: (  
    ([now],[their],[farm],[forever],[behind],[will not stay]),  
    ([τώρα],[τους],[αγρόκτημά], [για πάντα], [πίσω], [δεν θα παραμείνει]),  
  ),  
  line_styles: (text.with(fill:green),),  
  translation: "Now their farm will not stay behind forever.",  
)
```

(4)	Мы	с	Марко	поехали	автобусом	в	Перedelкино
	My	s	Marko	poexa-l-i	avtobus-om	v	Peredelkino
	1PL	COM	Marko	go-PST-PL	bus-INS	ALL	Peredelkino
	we	with	Marko	go-PST-PL	bus-by	to	Peredelkino
	nous	avec	Marco	allés	en bus	à	Peredelkino
	私たちは	と	マルコ	行きました	バスで	に	ペレデルキノ

Marko and I went to Peredelkino by bus

```
#numbered_gloss(  
  source_text: ([Мы], [с], [Марко], [поехали], [автобусом], [в], [Перedelкино]),  
  source_text_style: strong,  
  transliteration: ([My], [s], [Marko], [poexa-l-i], [avtobus-om], [v],  
[Peredelkino]),  
  morphemes: ([1#pl], [#com], [Marko], [go-#pst\-#pl], [bus-#ins], [#all],  
[Peredelkino]),  
  gloss_lines: (  
    ([we], [with], [Marko], [go-#pst\-#pl], [bus-by], [to], [Peredelkino]),  
    ([nous], [avec], [Marco], [allés], [en bus], [à], [Peredelkino]),  
    ([私たちは],[と],[マルコ],[行きました],[バスで],[に],[ペレデルキノ])  
  ),  
  line_styles: (  
    none,  
    text.with(style: "italic", fill: purple, font: "Noto Serif"),  
    text.with(font: "Noto Sans CJK JP", fill: red),  
  ),  
  translation: "Marko and I went to Peredelkino by bus",  
  translation_style: text.with(style: "italic", weight: "bold")  
)
```

3 Spacing between items

By default, horizontal space is placed between each item of a gloss. The amount of space is controlled by the `interword_spacing` parameter, which defaults to 1em but can be set to any desired length.

However, in some circumstances, one may want a particular item to not have an added space between it and one or more of the surrounding elements. This is controlled as part of per-item styling (see Section 8).

4 Line spacing

This package allows you to customize the spacing between the various lines in the gloss. The relevant spacing parameters are as follows:

gloss_line_spacing (Default: 0.5em) the amount of vertical space between lines of the gloss.

post_header_space (Default: 0.5em) the amount of extra vertical spacing added between the header text and the gloss itself. If set to `none`, a `linebreak()` will be used.

pre_translation_space (Default: 0.5em) the amount of extra vertical spacing added between the gloss itself and the translation text. If set to `none`, a `linebreak()` will be used.

While these parameters are set to sensible defaults, feel free to customize them as needed for your use case.

(5) Hittite (Lehmann 1982:211)

n=an apedani mehuni essandu.
CONN=him that.DAT.SG time.DAT.SG eat.they.shall
They shall celebrate him on that date

```
#numbered_gloss(  
  header_text: [Hittite (Lehmann 1982:211)],  
  source_text: ([n=an], [apedani], [mehuni],[essandu.]),  
  morphemes: ([#smallcaps[conn]=him], [that.#dat.#sg], [time.#dat.#sg],  
[eat.they.shall]),  
  translation: "They shall celebrate him on that date",  
  gloss_line_spacing: 1em,  
  interword_spacing: 3em,  
  post_header_space: 2em,  
  pre_translation_space: none,  
)
```

5 Gloss indentation

Oftentimes you may want to customize the amount of padding used to indent various parts of the gloss. The following parameters are provided to allow you to customize how much space is used for these elements.

total_padding (Default: 2.5em) the size of the padding from the left margin for the entire gloss example. This length marks where the text itself starts for both numbered and un-numbered glosses, allowing for consistent alignment between the two.

number_padding (Default: .5em) the size of the padding from the left margin for the number of a numbered gloss. Should be less than `gloss_padding` to avoid ugly overlaps. Is completely ignored in un-numbered glosses.

gloss_padding (Default: none) the size of the additional indentation for the aligned gloss itself. When set to none (as it is by default), the aligned gloss is not indented any further than the unaligned textual elements.

(6) Tagalog

b<um>ili
<ACTFOC>buy

“buy”

```
#numbered_gloss(  
  total_padding: 5em,  
  number_padding: 2em,  
  gloss_padding: 2em,  
  header_text: [Tagalog],  
  source_text: ([b<um>ili],),  
  morphemes: ([<#smallcaps([actfoc])>buy],),  
  translation: ["buy"]  
)
```

6 Item alignment

While the default spacing and alignment for items in glosses should be suitable for most use cases, there are cases in which you may want to customize these elements to achieve the desired result. `typst-leipzig-glossing` provides some added functionality for those who desire this level of control.

By default, all items in a gloss are aligned to the left of their respective stacks. This allows one to change the alignment of the gloss example as a whole without changing the alignment of the items within the aligned gloss itself:

(7) German

<i>unser-n</i>	<i>Väter-n</i>
our-DAT.PL	father\PL-dat

“to our fathers”

```
#align(center,  
  numbered_gloss(  
    header_text: [German],  
    source_text: ([unser-n], [Väter-n]),  
    morphemes: ([our-#dat.#pl],[father\\#pl\\-#dat.pl]),  
    translation: ["to our fathers"],  
  )  
)
```

The `item_alignment` parameter allows you to change this default alignment for all the aligned gloss lines if desired.

(8) Jaminjung (Schultze-Berndt 2000:92)

nanggayan guny-bi-yarluga?
who 2DU>3SG-FUT-poke

“Who do you two want to spear?”

```
#numbered_gloss(  
  item_alignment: center,  
  header_text: [Jaminjung (Schultze-Berndt 2000:92)],  
  source_text: ([nanggayan],[guny-bi-yarluga?]),  
  morphemes: ([who],[2#du>3#sg\~#fut\~poke]),  
  translation: ["Who do you two want to spear?"]  
)
```

To change the alignment of the individual lines, one can use the style parameters as usual. To prevent these style parameters from being overridden by `item_alignment` for the aligned gloss lines, you must set `item_alignment` to none.

(9)

Belhare

ne-e	a-khim-chi	n-yuNNa
DEM-LOC	1SG.POSS-house-PL	3NSG-be.NPST
DEM-LOC	1SPOSS-house-PL	3NS-be.NPST

“Here are my houses.”

```
#align(center,  
  numbered_gloss(  
    item_alignment: none,  
    header_text: [Belhare],  
    header_text_style: align.with(right),  
    source_text: ([ne-e],[a-khim-chi],[n-yuNNa]),  
    source_text_style: x => align(left, text(fill:red,x)),  
    morphemes: ([#dem\~#loc],[1#sg.#poss\~house-#pl],[3#smallcaps([nsg])-  
be.#smallcaps([npst]))],  
    gloss_lines: (  
      ([#dem\~#loc],[1#smallcaps([sposs])\~house-#pl],[3#smallcaps([ns])-  
be.#smallcaps([npst]))],  
    ),  
    line_styles: ((x => align(right, text(fill:blue,x))),),  
    translation: ["Here are my houses."],  
    translation_style: align.with(left)  
  )  
)
```

One may also align one particular item differently from the others. This is controlled as part of per-item styling (see Section 8).

7 Nlevel glossing: alternate syntax

The popular `expex` package for LaTeX provides an alternate syntax to use for glosses that they call “nlevel”. In this syntax, rather than writing the gloss line-by-line, one instead groups the words item-by-item, writing each word with its glosses for each line in turn. Advocates for this type of

syntax argue that it makes the resulting code more readable, since each word is grouped with its gloss, as well as easier to copy/paste recurring items from one gloss to another.

Nlevel glossing is *not* the default for this module, but those who prefer it can use it by setting the `nlevel` parameter of a gloss to `true`.

As an example, consider the following gloss:

(10) French

<i>aux</i>	<i>chevaux</i>
/o/	/ʃəvo/
to-ART-PL	horse.PL
to the	horses

to the horses

```
#numbered_gloss(  
  header_text: [French],  
  source_text: ([aux],[chevaux]),  
  transliteration: ([/o/],[/ʃəvo/]),  
  morphemes: ([to-#art\-#pl],[horse.#pl]),  
  gloss_lines: (  
    ([to the],[horses]),  
  ),  
  translation: "to the horses"  
)
```

An identical gloss to that produced by the default syntax above can be produced using `nlevel` syntax as follows:

```
#numbered_gloss(  
  nlevel: true,  
  header_text: [French],  
  gloss_lines: (  
    ([aux], [/o/], [to-#art\-#pl], [to the]),  
    ([chevaux], [/ʃəvo/], [horse.#pl], [horses]),  
  ),  
  line_styles: (emph,),  
  translation: "to the horses"  
)
```

Note that the `nlevel` example exclusively uses `gloss_lines` and `line_styles` for the aligned gloss text – since the `source_text`, `transliteration`, and `morphemes` parameters are based on a line-by-line syntax, they and their respective style parameters may not be used with `nlevel` syntax.

To avoid frustration for those not using `nlevel` syntax, the default behavior of italicizing the first line of the aligned gloss (corresponding to the `source_text` parameter in the default syntax) is not duplicated for `nlevel` glosses and must be manually specified. The behavior of all other parameters and defaults is unchanged and should be identical between both versions of the glossing syntax.

8 Per-item Styling

In certain cases, you may wish to format an individual item differently from the other items in the aligned gloss. To enable this, one can use the `item_styles` parameter. This parameter must be

formatted as a dictionary in which the keys are strings of the indices of the items and the values are the style function you would like applied to the item at that index in the aligned gloss. Because it's a dictionary, you only need to include items you want to have non-default styling – items whose indices aren't in this dictionary will be styled as default.

(11) German

Ich hab-e dein-e Mutter ge-seh-en
 I have-1SG 2PL.POSS-F mother PTCP-see-PTCP

"I saw your mother."

```
#numbered_gloss(
  header_text: [German],
  source_text: ([Ich],[hab-e],[dein-e],[Mutter],[ge-seh-en]),
  morphemes: ([I],[have-1#sg],[2#pl.#poss\-#f],[mother],[#ptcp\-see\-#ptcp]),
  translation: ["I saw your mother."],
  item_styles: (
    "2": text.with(fill:red),
    "4": text.with(fill:blue)
  ),
)
```

Sometimes you may not want a space before and/or after a particular item in the gloss. If you provide the string "no_space_before" or "no_space_after" as the value for an item in `item_styles`, this item will be flush against the item before or after it respectively, without any added horizontal space. This can be particularly useful for those who prefer to align their glosses morpheme-by-morpheme rather than word-by-word.

(12) Italian

and-iamo a -l supermercato
 go -PRS.1PL to-the supermarket

```
#numbered_gloss(
  header_text: [Italian],
  source_text: ([and],[-iamo],[a],[-l],[supermercato]),
  morphemes: ([go],[-#prs.1#pl],[to],[-the],[supermarket]),
  item_styles: ("1": "no_space_before", "2": "no_space_after")
)
```

Since use of spacing combined with right or left alignment for prefixes and suffixes is likely to be fairly common for this type of use-case, the string "prefix" can be used to indicate an item should be right-aligned with no space afterwards, and the string "suffix" can be used to indicate it should be left-aligned with no space beforehand.

(13) Tagalog

bi~bili
 IPFV~buy

"is buying"

```
#numbered_gloss(
  header_text: [Tagalog],
  source_text: ([bi~],[bili]),
  morphemes: ([#ipfv~],[buy]),
  translation: ["is buying"],
  item_styles: ("0": "prefix")
)
```

(14) Hebrew

yerak~rak -im
green~ATT-M.PL

“greenish ones”

```
#numbered_gloss(  
    header_text: [Hebrew],  
    source_text: ([yerak],[~rak],[-im]),  
    morphemes: ([green],[~#smallcaps([att])],[-  
#m.#pl]),  
    translation: ["greenish ones"],  
    item_alignment: right,  
    item_styles: ("1": "suffix", "2": "suffix"),  
)
```

If you would like to apply both spacing and other styling to a given item beyond these particular combinations, you should provide a dictionary as the value in `item_styles`. This dictionary should contain the following three keys:

- `style`: whatever style function you would like applied to the given item
- `spacing_before`: Defaults to true. Set to false if you would not like spacing before this item.
- `spacing_after`: Defaults to true. Set to false if you would not like spacing after this item.

(15) Mandarin

猫 可以 吃 点 香肠 作为 犒赏
cat may eat little.bit fragrant-intestines do-as reward

“Cats may have a little sausage as a treat.”

```
#numbered_gloss(  
    header_text: [Mandarin],  
    source_text: ([猫],[可以],[吃],[点],[香],[肠],[作],[为],[犒赏]),  
    morphemes: ([cat],[may],[eat],[little.bit],[fragrant],[-intestines],[do],[-  
as],[reward]),  
    item_styles: (  
        "0": text.with(fill:red),  
        "1": x => align(right,text(fill: green, x)),  
        "2": x => align(left,text(fill: green, x)),  
        "3": x => align(center,text(fill: purple, x)),  
        "4": (  
            "style": x => align(right, text(fill: orange,x)),  
            "spacing_after": false,  
        ),  
        "5": text.with(fill:orange),  
        "6": text.with(fill:blue),  
        "7": (  
            "style": x => align(left,text(fill: blue, x)),  
            "spacing_before": false,  
        ),  
        "8": x => align(center,text(fill: green,x))  
    ),  
    translation: ["Cats may have a little sausage as a treat."]  
)
```

Nlevel-specific options