The covington Package Macros for Linguistics

Michael A. Covington Jürgen Spitzmüller*

Version 1.6, September 7, 2018

Abstract

This package, initially a collection of Michael A. Covington's private macros, provides numerous minor LaTeX enhancements for linguistics, including multiple diacritics on the same letter, interlinear glosses (word-by-word translations), Discourse Representation Structures, and example numbering.

Contents

1	Introduction	2
2	Stacked diacritics	2
3	Example numbers	3
4	The example environment	4
5	The examples environment	5
6	The subexamples environment	5
7	Referring to examples	6
8	Glossing sentences word-by-word	7
9	Phrase structure rules	9
10	Feature structures	9
11	Discourse Representation Structures	10
12	Exercises	11
13	Reference Lists	12
14	Displayed sentences	13
15	Big curly brackets (disjunctions)	13
16	Release history	13

^{*}Current maintainer. Please report issues via https://github.com/jspitz/covington

1 Introduction

This is the documentation for version 1.6 of covington (September 7, 2018), which is a LATEX package providing macros for typing some special notations common in linguistics.¹

To use covington with $ETEX a_{\mathcal{E}}$, load the package as usual by adding the command $\scale{covington}$ to your document preamble. The package has the following options:

force: Force the redefinition of environments that have already been defined by other packages or the class.

This applies to the **example**, **examples**, **subexamples** and **exercise** environments, which are by default not touched if they are already defined before covington is loaded. See sec. 4, 5, 6 and 12 for details.

keeplayout: Do not tweak the layout.

Covington sets \raggedbottom and redefines the value of the \textfloatsep length. This just follows the preferences of the original package author and is not necessary for the package's functionality. Yet for backwards compatibility reasons, we cannot change this. Thus, we provide the option described here to opt out this presetting.

In what follows we presume that you know how to use ET_EX and have access to ET_EX manuals. Note that covington does not provide any special fonts or character sets. However, it can be used in combination with other style sheets that do.

If you are using covington and uga.sty (UGa thesis style) together, you should load uga before covington.

2 Stacked diacritics

LATEX provides a generous range of diacritics that can be placed on or below any letter, such as:

 $\grave{\textbf{x}}\mathrel{\acute{\textbf{x}}} \grave{\boldsymbol{x}} \mathrel{\ddot{\textbf{x}}} \check{\boldsymbol{x}} \mathrel{\ddot{\textbf{x}}} \check{\boldsymbol{x}} \mathrel{\dot{\textbf{x}}} \grave{\boldsymbol{x}} \mathrel{\dot{\textbf{x}}} \grave{\boldsymbol{x}} \mathrel{\dot{\textbf{x}}} \grave{\boldsymbol{x}} \mathrel{\dot{\textbf{x}}} \grave{\boldsymbol{x}}$

which are typed, respectively, as:

```
'{x} ''{x} ''{x} '''{x} '''{x} '={x} 'H{x} 't{xx} 'c{x} 'b{x}
```

 ${
m I\!M}_{
m E\!X}$ also provides support for many non-ASCII characters, such as²:

ı jæÆœŒåÅøØłŁß;;

¹The package has a long history. It started off as a collection of private macros back in the $\[mathscale{LTEX}\]$ 2.09 days and was initially released as covingtn.sty (following the old 8.3 FAT file name limit). In emTEX under Ms-Dos, the file was distributed as covingto.sty. Eventually, it has been renamed to covington and adapted to $\[mathscale{LTEX}\]$ 2 $_{\mathcal{E}}$.

²Please refer to [1] for a comprehensive list of special characters and symbols.

via the macros:

i j AE OE AE OE AA O O L L SS ' '

Out of the box, however, LATEX doesn't give you a convenient way to put *two* diacritical marks on the same letter. To fill this gap, covington provides the following macros:

```
\text{wodias{~upper diac.>}{<lower diac.>}{<char>} to combine any two diacritics, e. g., \text{wodias}{-}{} = \tilde{\bar{a}}
```

```
\acm{...}for acute over macron, e. g., \acm{a} = \dot{\bar{a}}\grm{...}for grave over macron, e. g., \grm{a} = \dot{\bar{a}}\cim{...}for circumflex over macron, e. g., \cim{a} = \hat{\bar{a}}
```

The first of these is the general case³ and the latter three are special cases that are often used in Greek transcription. Now you can type $Koin\dot{e}$ with both accents in place.

The vertical distance between the two diacritics can be adjusted via the macro **\SetDiaOffset{<length>}** which lets you increase or decrease the vertical space that is currently in effect. If you'd use **\SetDiaOffset{-0.25ex}**, the above examples would come out as

```
\twodias{<upre diac.>}{<lower diac.>}{<char>} to combine any two diacritics, e. g., \twodias{\~}{=}{a} = \tilde{\bar{a}}
```

```
\acm{...}for acute over macron, e. g., \acm{a} = \dot{a}\grm{...}for grave over macron, e. g., \grm{a} = \dot{a}\cim{...}for circumflex over macron, e. g., \cim{a} = \dot{a}
```

with a slightly better matching distance for the font used here.

Note that not all accent macros work in the tabbing environment. Use the Tabbing package or refer to [1] for alternative solutions.

3 Example numbers

Linguistics papers often include numbered examples. The macro **\exampleno** generates a new example number and can be used anywhere you want the number to appear. For example, to display a sentence with a number at the extreme right, do this:

```
\begin{flushleft}
This is a sentence. \hfill (\exampleno)
\end{flushleft}
```

Here's what you get:

This is a sentence.

(1)

³Alternatively, there's also the old syntax \twoacc[<upper diac.>|<char with lower diacr.>], e.g. \twoacc[\-|\={a}] to the same effect, which is however discouraged due to its rather odd form.

The example counter is actually the same as LATEX's equation counter, so that if you use equations and numbered examples in the same paper, you get a single continuous series of numbers. If you want to access the number without changing it, use \theequation.

Also, you can use \label and \ref with example numbers in exactly the same way as with equation numbers (see sec. 7 for details). This applies to the **example**, **examples** and **subexamples** environments, described next, as well as to **exampleno** itself.

4 The example environment

The **example** environment (alias **covexample**) displays a single example with a generated example number to the left of it. If you type

```
\begin{example}
This is a sentence.
\end{example}
```

or

```
\begin{covexample}
This is a sentence.
\end{covexample}
```

you get:

(2) This is a sentence.

The example can be of any length; it can consist of many lines (separated by $\)$, or even whole paragraphs.

If you need more space between the example number and the text, you can increase it by means of the length \examplenumbersep (which is preset to Opt). Doing \setlength\examplenumbersep{1em}, for instance, will increase the space by 1 em.

Note that, as of version 1.1, covington checks if there is already an **example** environment defined (e.g., by the class). If so, covington does not define its own one. However, there is always the alias environment **covexample** which can be used in order to produce covington's example. If you use the package option **force**, covington will override existing **example** environments. In any case, the package will issue a warning if **example** is already defined (this is the case, for instance, if you use covington with the beamer class).

One way to number sub-examples is to use itemize or enumerate within an example, like this:

```
\begin{example}
\begin{itemize}
\item[(a)] This is the first sentence.
\item[(b)] This is the second sentence.
\end{itemize}
\end{example}
```

This prints as:

(3) (a) This is the first sentence.

(b) This is the second sentence.

However, the **examples** and **subexamples** environments, described in turn, are usually more convenient.

5 The examples environment

To display a series of examples together, each with its own example number, use **examples** (or **covexamples**) instead of **example** or **covexample**. The only difference is that there can be more than one example, and each of them has to be introduced by \item, like this:

```
\begin{examples}
\item This is the first sentence.
\item This is the second sentence.
\end{examples}
```

or, respectively:

```
\begin{covexamples}
\item This is the first sentence.
\item This is the second sentence.
\end{covexamples}
```

This prints as:

- (4) This is the first sentence.
- (5) This is the second sentence.

As for example, covington checks if there is already an examples environment defined, and if this is the case, covington does not define its own one. The alias environment covexamples is always available as a fallback. If you use the package option force, covington will override existing examples environments. The package will issue a warning if examples is already defined (this is the case, for instance, if you use covington with the beamer class), telling you how it has dealt with the situation.

6 The subexamples environment

Sometimes a set of (paradigmatic) sub-examples gets only one main example number with alphabetic sub-numbering, as in (6 a). To achieve this most conveniently, covington provides the **subexamples** (or **covsubexamples**) environment. The difference to **examples**/**covexamples** is the numbering:

```
\begin{subexamples}
\item This is the first sentence.
\item This is the second sentence.
\end{subexamples}
```

or, respectively:

```
\begin{covsubexamples}
\item This is the first sentence.
\item This is the second sentence.
\end{covsubexamples}
```

prints as:

- (6) (a) This is the first sentence.
 - (b) This is the second sentence.

Again, covington checks if there is already an **subexamples** environment defined, and if this is the case, covington does not define its own one. The alias environment **covsubexamples** is always available as a fallback. If you use the package option **force**, covington will override existing **subexamples** environments. The package will issue a warning if **subexamples** is already defined.

If you do not want the subexample letter in parentheses, you can redefine (via \renewcommand*) the macro \covsubexnumber which has the following default definition:

```
\newcommand*\covsubexnumber[1]{(#1)}
```

The distance between example number and subnumber (letter) can be changed via the length **\examplenumbersep** (which is preset to Opt). The distance between example subnumber and text can be changed via the length **\subexamplenumbersep** (preset to Opt as well). Doing

```
\setlength{\examplenumbersep}{-0.2em}
\setlength{\subexamplenumbersep}{0.5em}
```

for instance, will come out like this:4

- (7) (a) This is the first sentence.
 - (b) This is the second sentence.

7 Referring to examples

References to examples and sub-examples can be made the usual way via the **\ref** command. The references do not have parentheses by default, i. e., a reference to the example in section 4 would be printed as 2, a reference to the sub-example in section 6 as 6 a. For convenience, covington provides a command **\pxref** that also prints the parentheses, as in (2) and (6 a).

⁴Note, though, that negative shrinking is only possible to a certain degree in the first subitem.

8 Glossing sentences word-by-word

To gloss a sentence is to annotate it word-by-word. Most commonly, a sentence in a foreign language is followed by a word-for-word translation (with the words lined up vertically) and then a smooth translation (not lined up), like this:⁵

Dit is een Nederlands voorbeeld. This is a Dutch example.

'This is an example in Dutch.'

That particular example would be typed as:

```
\gll Dit is een Nederlands voorbeeld.
This is a Dutch example.
\glt 'This is an example in Dutch.'
\glend
```

Notice that the words do not have to be typed lining up; instead, T_EX counts them. If the words in the two languages do not correspond one-to-one, you can use curly brackets to group words. For example, to print

Dit is een voorbeeldje in het Nederlands. This is a little example in Dutch. 'This is a little example in Dutch.'

you would type:

```
\gll Dit is een voorbeeldje in het Nederlands.
This is a {little example} in {} Dutch.
\glt 'This is a little example in Dutch.'
\glend
```

Note that covington locally activates the end of line in glosses in order to identify the different lines of the gloss (via category code change). This does not work inside macros (e. g., if the gloss is in a footnote). To work around this, a special version of the \gll macro is provided that does without the character activation: \xgll. This can also be used in macro arguments; however, the end of each gloss line needs to be explicitly specified by the \xgle macro in this case. If you want to put the above gloss in a footnote, thus, you would type:

```
\xgll Dit is een voorbeeldje in het Nederlands.\xgle
This is a {little example} in {} Dutch.\xgle
\glt 'This is a little example in Dutch.'
\glend
```

All together, covington provides eight macros for dealing with glosses:

 $^{^5\}mathrm{The}$ macros for handling glosses are adapted with permission from gloss.tex, by Marcel R. van der Goot.

- \gll introduces two lines of words vertically aligned, and activates an environment very similar to flushleft. The two lines are separated by a normal line break (carriage return).
- \gll is like \gll except that it introduces *three* lines of lined-up words (useful for cited forms, morphology, and translation).
- \xgll is similar to \gll except that it does not make the line ending active. It thus works inside macros such as footnotes but requires explicit gloss line termination via \xgle.
- \xglll is similar to \glll except that it does not make the line ending active. It thus works inside macros such as footnotes but requires explicit gloss line termination via \xgle.
- \xgle is a gloss line ending marker to be used with \xgll and \xglll.
- \glt ends the set of lined-up lines and introduces a line (or more) of translation.
- \gln is like \glt but does not start a new line (useful when no translation follows but you want to put a number on the right).
- \glend ends the special flushleft-like environment.

Here are several examples. First, a sentence with three lines aligned, instead of just two:

Hocestaliudexemplum.n.sg.nom3.sgn.sg.nomn.sg.nomThisisanotherexample.

'This is another example.'

This is typed as:

```
\gll Hoc est aliud exemplum.
    n.sg.nom 3.sg n.sg.nom n.sg.nom
    This is another example.
\glt 'This is another example.'
\glend
```

Next, an example with a gloss but no translation, with an example number at the right:

(8)

Hoc habet numerum. This has number

That one was typed as:

```
\gll Hoc habet numerum.
This has number
\gln \hfill (\exampleno)
\glend
```

Finally we'll put a glossed sentence inside the example environment, which is a very common way of using it:

 (9) Hoc habet numerum praepositum. This has number preposed
 'This one has a number in front of it.'

This last example was, of course, typed as:

```
\begin{example}
\gll Hoc habet numerum praepositum.
    This has number preposed
\glt 'This one has a number in front of it.'
\glend
\end{example}
```

Notice that every glossed sentence begins with either \gll or \gll, then contains either \glt or \gln, and ends with \glend. Layout is critical in the part preceding \glt or \gln, and fairly free afterward.

9 Phrase structure rules

To print phrase structure rules such as $S \rightarrow NP VP$ you can use covington's macro $\psr{<constituent>}{(sub-constituents>)} (for the given example, <math>\psr{S}{NP-VP})$.

10 Feature structures

To print a feature structure such as

```
case : nom person : P
```

you can type:

```
\fs{case:nom \\ person:P}
```

The feature structure can appear anywhere — in continuous text, in a displayed environment such as flushleft, or inside a phrase-structure rule, or even inside another feature structure.

To put a category label at the top of the feature structure, like this,

```
N

[ case : nom

person : P
```

here's what you type:

\lfs{N}{case:nom \\ person:P}

And here is an example of a PS-rule made of labeled feature structures:

S	\rightarrow	NP	VP
[tense:T]		case : nom	tense : T
		number : N	number : N

which was obviously coded as:

\psr{\lfs{S}{tense:T}}
 {\lfs{NP}{case:nom \\ number:N}
 \lfs{VP}{tense:T \\ number:N} }

11 Discourse Representation Structures

Several macros in covington facilitate display of discourse Representation Structures (DRses) in the box notation introduced by Hans Kamp. The simplest one is \drs, which takes two arguments: a list of discourse variables joined by ~, and a list of DRs conditions separated by \\. Nesting is permitted. Note that the \drs macro itself does not give you a displayed environment; you must use flushleft or the like to display the DRs. Here are some examples:



To display a sentence above the DRS, use \sdrs, as in:



which prints as:

A donkey is green.



Some DRS connectives are also provided (normally for forming DRSes that are to be nested within other DRSes). The macro **\negdrs** forms a DRS preceded by a negation symbol:

\negdrs{X}{donkey(X)\\green(X)}



Finally, \ifdrs forms a pair of DRSes joined by a big arrow, like this:

```
\ifdrs{X}{donkey(X)\\hungry(X)}
        {~}{feeds(Pedro,X)}
```



If you have an "if"-structure appearing among ordinary predicates inside a DRS, you may prefer to use **\alifdrs**, which is just like **\ifdrs** but shifted slightly to the left for better alignment.

12 Exercises

The **exercise** environment (alias **covexercise**) generates an exercise numbered according to chapter, section, and subsection (suitable for use in a large book; in this example, the subsection number is going to come out as o). Here is an example:

Exercise 12.0.1 (Project) Prove that the above assertion is true.

This was coded as

```
\begin{exercise}[Project]
Prove that the above assertion is true.
\end{exercise}
```

The argument ([Project] in the example) is optional.

Note that, as of version 1.1, covington checks if there is already an **exercise** environment defined (e. g., by the class). If so, covington does not define its own one. However, there is always the alias environment **covexercise** which can be used in order to produce covington's exercise. If you use the package option **force**, covington will override existing **exercise** environments. In any case, the package will issue a warning if **exercise** is already defined.

13 Reference Lists

To type a simple LSA-style hanging-indented reference list, you can use the **reflist** environment. (*Note:* **reflist** is not integrated with BibTEX in any way.⁶) For example,

```
\begin{reflist}
Barton, G. Edward; Berwick, Robert C.; and Ristad, Eric Sven. 1987.
Computational complexity and natural language. Cambridge,
Massachusetts: MIT Press.
Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge,
Massachusetts: MIT Press.
Covington, Michael. 1993. Natural language processing for Prolog
programmers. Englewood Cliffs, New Jersey: Prentice-Hall.
\end{reflist}
```

prints as:

- Barton, G. Edward; Berwick, Robert C.; and Ristad, Eric Sven. 1987. Computational complexity and natural language. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, Massachusetts: MIT Press.
- Covington, Michael A. 1993. Natural language processing for Prolog programmers. Englewood Cliffs, New Jersey: Prentice-Hall.

By default, the references have a hanging indentation of 3 em. This can be globally changed by altering the length \reflistindent. Doing \setlength\reflistindent{1.5em}, for instance, will shorten the indentation by half. Likewise, the length \reflistitemsep (6 pt by default) and \reflistparsep (ca. 4 pt by default) can be adjusted to alter the vertical separation (\itemsep and \parsep, for that matter) of reference entries.

Notice that within the reference list, "French spacing" is in effect — that is, spaces after periods are no wider than normal spaces. Thus you do not have to do anything special to avoid excessive space after people's initials.

⁶For BibTEX, there are several options: the LSA style, as used in the journal *Language*, can be obtained by means of the style files lsalike.bst (http://www.icsi.berkeley.edu/ftp/pub/speech/jurafsky/lsalike. bst) or language.bst (http://ron.artstein.org/resources/language.bst); the latter uses natbib. The socalled *Unified Style Sheet for Linguistics*, as proposed by the CELXJ (*Committee of Editors of Linguistics Journals*), which slightly differs from the LSA style, is followed by the style file unified.bst (available at http://celxj.org/downloads/unified.bst). A biblatex style file for the unified style is available at https://github.com/semprag/biblatex-sp-unified.

14 Displayed sentences

The macro \sentence displays an italicized sentence (it is a combination of flushleft and itshape). If you type

```
\sentence{This is a sentence.}
```

you get:

This is a sentence.

15 Big curly brackets (disjunctions)

Last of all, the two-argument macro **\either** expresses alternatives within a sentence or PS-rule:

the \either{big}{large} dog = the
$$\begin{cases} big \\ large \end{cases}$$
 dog
\psr{A}{B-\either{C}D}-E} = $A \rightarrow B \begin{cases} C \\ D \end{cases} E$

That's all there is for now. Suggestions for improving covington are welcome, and bug reports are actively solicited (via https://github.com/jspitz/covington). Please note, however, that this is free software, and the authors make no commitment to do any further work on it.

16 Release history

1.6 (2018 September 7)

- Introduce new environment subexamples (see sec. 6).
- Introduce new command \pxref (see sec. 7).

1.5 (2018 August 24)

• Introduce new option **keeplayout** which allows to opt-out the layout presettings covington does (\raggedbottom, \textfloatsep).

1.4 (2017 May 23)

- Introduce a new macro \twodias that supersedes the rather odd \twoacc (which is kept for backwards compatibility). See sec. 2 for details.
- Introduce macro **\SetDiaOffset** for more convenient setting of vertical distance in stacked diacritics. See sec. 2 for details.
- LATEX 2.09 is no longer officially supported (it might continue to work, but is not tested).

1.3 (2017 April 5)

- Gloss variants \xgll and \xgll that work inside macros (such as footnotes) but require explicit gloss line end markers (\xgle). See sec. 8 for details.
- New lengths \reflistindent, \reflistparsep and \reflistitemsep to globally adjust the indentation or vertical space, respectively, of reflist items. See sec. 13 for details.

1.2 (2016 August 26)

- New length \examplenumbersep to adjust (increase) the horizontal space between example number and example text. See sec. 4 for details.
- Add some more info about bibliography generation.

1.1a (2016 July 7)

• Fix encoding problem in documentation and some typos. No change in functionality.

1.1 (2016 July 6)

- The package now uses NFSS font commands if available (fallback for LATEX 2.09 is still provided).
- Work around clash with classes/packages that define their own example and examples environments (most notably the beamer class) as well as execise environments. The covington package no longer blindly attempts to define these environments. By default, it does not define them if they are already defined (covington's own environments, however, are still available via aliases). By means of a new package option, a redefinition can also be forced. See sec. 4 and 5 for details.
- New length \twoaccsep allows for the adjustment of the distance between stacked accents (see sec. 2).
- Update manual.
- New maintainer: J. Spitzmüller.
- License has been changed to LPPL (in agreement with M. Covington)
- Introduce version numbers. Arbitrarily, we start with 1.1.

2014 May 16

- Patches by Robin Fairbairns:
 - Setting of \textfloatsep uses \setlength rather than \renewcommand
 - Style file converted to un*x line endings

2001 March 27

- It is no longer necessary to type \it to get proper italic type in feature structures.
- Instructions have been rewritten with ${\rm I\!A} T_{\!E\!} X \, {\rm 2}_{\mathcal E}$ users in mind.

Older versions

- Multiple accents on a single letter (e. g., $\dot{\tilde{a}})$ are supported.
- This package is now called covington (with the o) and is compatible with $\[MTeX 2_{\mathcal{E}}\]$ and NFSS as well as $\[MTeX 2.09\]$.
- The vertical placement of labeled feature structures has been changed so that the category labels line up regardless of the size of the structures.

References

 Pakin, Scott. The Comprehensive LATEX Symbol List. 30 November 2015. http: //www.ctan.org/pkg/comprehensive.