The babel-kazakh package

Kazakh language support for pdfLaTeX, XeLaTeX, and LuaLaTeX

Berik I. Tuleuov

Astana, Kazakhstan

berik_t@yahoo.com

Documentation revised November 21, 2026

Abstract

The babel-kazakh package provides Kazakh language support for LaTeX. It supports the traditional Babel interface with pdfLaTeX, XeLaTeX, and LuaLaTeX, and it also provides a Polyglossia language definition for the Unicode engines. The package supplies Kazakh captions, date strings, typographical conventions, and preliminary BibLaTeX localization. Kazakh locale data for the modern Babel interface is maintained in the main Babel distribution.

Contents

1 Background and scope **2**

2 Package components **2**

3 Engine and package support **3**

4 Quick start with Babel **3**

4.1 pdfLaTeX 3

4.2 XeLaTeX and LuaLaTeX 3

5 Quick start with Polyglossia **4**

6 Kazakh numbering and captions **5**

7 Bibliography support **5**

8 Hyphenation tests **6**

9 The supplied test suite **7**

9.1 Compiling the Babel driver 7

9.2 Compiling the Polyglossia driver 7

10 What the shared body verifies **8**

11 Installation notes 8

12 License and maintenance 9

1 Background and scope

Work on the package followed the development of a Kazakh hyphenation dictionary in February 2026. The goal of babel-kazakh is to make Kazakh documents usable with all three major LaTeX engines while keeping the language conventions consistent between Babel, Polyglossia, and BibLaTeX.

The current support covers:

- Kazakh captions and Gregorian date names;
- the number–name order required for parts, chapters, appendices, figures, and tables;
- language switching in multilingual documents;
- Kazakh hyphenation when the corresponding patterns are installed;
- preliminary bibliography strings for BibLaTeX.

2 Package components

The runtime files supplied by babel-kazakh are listed below.

File	Purpose
`kazakh.ldf`	Traditional Babel language definition.
`kazakh.sty`	Compatibility interface for Kazakh support.
`gloss-kazakh.ldf`	Kazakh language definition for Polyglossia.
`gloss-kk.ldf`	BCP 47 alias used by Polyglossia.
`kazakh.lbx`	Kazakh localization strings for BibLaTeX.

The file `babel-kk.ini` is *not* distributed by this package. Starting with Babel 26.16 (November 10, 2026), the Kazakh locale file is part of the main Babel distribution and is installed as

`tex/generic/babel/locale/kk/babel-kk.ini`.

Corrections to that locale should therefore be reported to the main Babel project at https://github.com/latex3/babel/issues.

3 Engine and package support

Engine	Babel	Polyglossia	Bibliography test
pdfLaTeX	yes	no	BibLaTeX+Biber
XeLaTeX	yes	yes	BibLaTeX+Biber
LuaLaTeX	yes	yes	BibLaTeX+Biber

For pdfLaTeX, use UTF-8 input and the T2A font encoding. For the two Unicode engines, select OpenType fonts containing the Kazakh Cyrillic characters. Every main, sans-serif, or monospaced family used in Kazakh text must provide the required Cyrillic glyphs.

The package and the supplied tests were developed and tested with TeX Live 2025 under Linux. Under Windows, XeLaTeX and LuaLaTeX can use either fonts installed in Windows or suitable fonts installed with the TeX distribution. The test drivers name the DejaVu Serif, DejaVu Sans, and DejaVu Sans Mono families explicitly. If they are unavailable, install the DejaVu fonts or replace all three names with Windows fonts that contain the Kazakh Cyrillic characters. The main, sans-serif, and monospaced selections must all provide Cyrillic coverage; this is checked particularly strictly by Polyglossia. The pdfLaTeX test does not use Windows OpenType fonts: it relies on the T2A-encoded fonts installed by the TeX distribution.

4 Quick start with Babel

4.1 pdfLaTeX

The traditional interface uses `kazakh.ldf`:

```
\documentclass{article}
\usepackage[T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[kazakh]{babel}

\begin{document}
Қазақ тіліндегі мәтін.
\end{document}
```

Compile the document with:

```
pdflatex document.tex
```

4.2 XeLaTeX and LuaLaTeX

The modern locale interface is recommended for Unicode engines:

```
\documentclass{article}
\usepackage[english]{babel}
\usepackage{fontspec}
\setmainfont{DejaVu Serif}
\babelprovide[import,main]{kazakh}

\begin{document}
Қазақ тіліндегі мәтін.

\foreignlanguage{english}{An English sentence.}
\end{document}
```

Compile with either:

```
xelatex document.tex
```

or:

```
lualatex document.tex
```

5 Quick start with Polyglossia

Polyglossia requires XeLaTeX or LuaLaTeX. The test driver uses fonts from the DejaVu family because all three selected families contain Cyrillic glyphs.

```
\documentclass{article}
\usepackage{fontspec}
\setmainfont{DejaVu Serif}
\setsansfont{DejaVu Sans}
\setmonofont{DejaVu Sans Mono}
\usepackage{polyglossia}
\setmainlanguage{kazakh}
\setotherlanguage{english}

\begin{document}
Қазақ тіліндегі мәтін.

\foreignlanguage{english}{An English sentence.}
\end{document}
```

If a selected font lacks the Cyrillic script, Polyglossia reports an error. Set every font family actually used by the document to a font with Kazakh Cyrillic coverage.

6 Kazakh numbering and captions

Kazakh places the number before the common noun in the names tested by the package. The common noun begins with a lowercase letter:

Document element	Expected form
Part	**1 бөлім**
Chapter	**3 тарау**
Appendix	**А қосымша**
Figure	**7 сурет**
Table	**10 кесте**

The same ordering is used by the Unicode locale data, the LICR data needed by pdfLaTeX, the traditional `kazakh.ldf` implementation, and the Polyglossia definition.

The shared test body contains numbered parts, chapters, sections, an appendix, two tables, and two figures. For example:

```
\part{Қазақ тілін қолдау құралдарын тексеру}

\ifcsname chapter\endcsname
  \chapter{Тараулар мен қосымшаларды тексеру}
\fi

\begin{table}[htbp]
  \centering
  \caption{Қарапайым сандық кесте}
  ...
\end{table}

\begin{figure}[htbp]
  \centering
  ...
  \caption{Сынық сызықтың шартты графигі}
\end{figure}
```

Changing the class in either test driver from `article` to `report` or `book` activates the chapter tests. The `book` class also makes it possible to inspect the Kazakh order in running heads.

7 Bibliography support

The package contains `kazakh.lbx`, which provides Kazakh localization for BibLaTeX. This localization is still under development. The test database `refs.bib` contains Kazakh and English sources, including books, a dissertation, and an online resource.

A parent document loads the database as follows:

```
\usepackage[backend=biber]{biblatex}
\addbibresource{refs.bib}
```

Entries written in Kazakh use `langid = {kazakh}`. One example from the supplied database is:

```
@book{kz_orph,
  author    = {Сыздық, Р. and Уәлиұлы, Н. and
               Күдеринова, Қ. and Фазылжанова, А. and others},
  title     = {Қазақ тілінің орфографиялық сөздігі: 53 000 сөз},
  location  = {Алматы},
  publisher = {Арыс},
  date      = {2005},
  pagetotal = {616},
  langid    = {kazakh}
}
```

The complete compilation sequence is:

```
<engine> test-file.tex
biber test-file
<engine> test-file.tex
<engine> test-file.tex
```

Here `<engine>` is `pdflatex`, `xelatex`, or `lualatex`, as permitted by the selected language package.

8 Hyphenation tests

The Babel and Polyglossia drivers use the same list of Kazakh words. Two complementary mechanisms are employed:

- `\showhyphens` writes the possible break points to the log;
- the testhyphens environment prints the same kind of test material in the document, arranged in three columns.

The tests explicitly set

```
\lefthyphenmin=2
\righthyphenmin=2
```

and include native words, derived forms, loanwords, and difficult vowel sequences. Examples include *құмырсқа*, *күңгірттеу*, *жаңғыртты*, *балуан*, *тәрбие*, and *офтальмология*.

9 The supplied test suite

There are exactly two standalone, compilable test drivers. Both include the same headless body and use the same bibliography database.

File	Role
`kazakh-unversal-bib-test.tex`	Universal Babel driver. It selects pdfLaTeX settings automatically and uses the modern locale interface under XeLaTeX or LuaLaTeX.
`kazakh-polyglossia-bib-test.tex`	Polyglossia driver for XeLaTeX and LuaLaTeX. Under pdfLaTeX it produces an explanatory page instead of attempting to load Polyglossia.

The supporting files are:

`kazakh-test-body.tex`
: A headless fragment containing structural tests, formulas, cross references, tables, figures created with the standard LaTeX `picture` environment, an appendix, citations, and the bibliography command.

`variables.sty`
: A local helper belonging to the test suite. It reports the engine, language package and version, document class, and class options in the generated test document. It is not a runtime component of babel-kazakh; keep it in the same directory as the two test drivers.

`refs.bib`
: Bibliographic data used by both drivers.

Both drivers begin with alternative `\documentclass` lines for the standard `article`, `book`, and `report` classes. To test another class, comment out the currently active line and uncomment the required one, leaving exactly one `\documentclass` command active. The shared body adapts to the selected class: chapter tests are performed only when the class provides chapters. The diagnostics generated by `variables.sty` record the selected class and its options.

9.1 Compiling the Babel driver

For pdfLaTeX:

```
pdflatex kazakh-unversal-bib-test.tex
biber kazakh-unversal-bib-test
pdflatex kazakh-unversal-bib-test.tex
pdflatex kazakh-unversal-bib-test.tex
```

For a Unicode engine, replace `pdflatex` by `xelatex` or `lualatex`.

9.2 Compiling the Polyglossia driver

With XeLaTeX:

```
xelatex kazakh-polyglossia-bib-test.tex
biber kazakh-polyglossia-bib-test
xelatex kazakh-polyglossia-bib-test.tex
xelatex kazakh-polyglossia-bib-test.tex
```

The same sequence works with LuaLaTeX after replacing `xelatex` by `lualatex`.

10 What the shared body verifies

The file `kazakh-test-body.tex` is deliberately independent of the language package and engine. Its checks are summarized below.

Area	Verification
Document structure	Parts, chapters when available, sections, subsections, subsubsections, and appendices.
Mathematics	A quadratic formula, a finite sum, equation numbering, and cross-references.
Tables	Two captioned tables containing numbers and Kazakh letters.
Figures	Two captioned drawings made entirely in LaTeX.
Bibliography	Kazakh and English citations plus a bibliography in the table of contents.
Diagnostics	Engine, package, package version, class, and class options.

Because the same body is included by both drivers, differences in captions, numbering, language switching, or bibliography strings can be compared directly.

11 Installation notes

In a current TeX Live installation the package files are placed in their normal TDS locations. When testing a development copy in a personal tree, run the filename-database update command required by the TeX distribution and verify the selected files with `kpsewhich`. For example:

```
kpsewhich kazakh.ldf
kpsewhich gloss-kazakh.ldf
kpsewhich gloss-kk.ldf
kpsewhich kazakh.lbx
kpsewhich babel-kk.ini
```

The final command should resolve to the main Babel tree, not to the babel-kazakh package directory.

12 License and maintenance

Copyright © 2026 Berik I. Tuleuov.

The babel-kazakh package may be distributed and/or modified under the conditions of the LaTeX Project Public License, version 1.3c or any later version. The work has the LPPL maintenance status `maintained`; the Current Maintainer is Berik I. Tuleuov.