

1 T_EX distributions

1.1 T_EX Live

In T_EX Live you can install different language collections `collection-lang<foo>`.

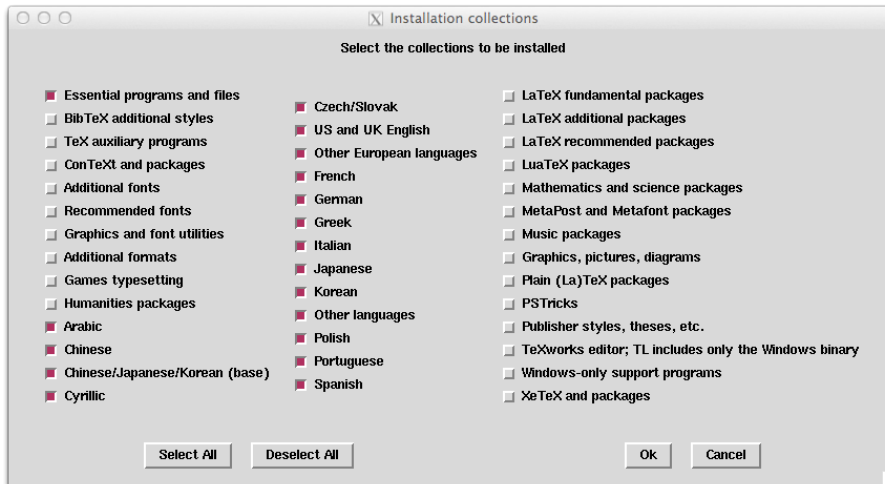


Figure 1

For example, the German collection `collection-langgerman` contains packages `hyphen-german` and `dehyph-expt1` which both contain hyphenation patterns (in addition to many other packages related to German language). The Greek collection `collection-langgreek` contains `hyphen-greek` and `hyphen-ancientgreek` among other packages. The package `dehyph-expt1` is maintained independently, while all packages `hyphen-<foo>` belong to `hyph-utf8`.

Hyphenation packages may contain one or more patterns. For example `hyphen-norwegian` contains two directives for adding two hyphenation patterns:

```
execute AddHyphen \
  name=bokmal synonyms=norwegian,norsk \
  lefthyphenmin=2 \
  righthyphenmin=2 \
  file=loadhyph-nb.tex \
  file_patterns=hyph-nb.pat.txt \
  file_exceptions=hyph-nb.hyp.txt
execute AddHyphen \
  name=nynorsk \
  lefthyphenmin=2 \
  righthyphenmin=2 \
  file=loadhyph-nn.tex \
  file_patterns=hyph-nn.pat.txt \
  file_exceptions=hyph-nn.hyp.txt
```

The keywords **name** and **synonyms** are used as language name to access the patterns in T_EX, **file** is used as pattern loader for 8-bit engines and X_YT_EX, while **file_patterns** and **file_exceptions** are plain text files used by LuaT_EX.

The information from `AddHyphen` directives end up in three different files: `language.dat` used by L^AT_EX, `language.def` used by plain ϵ -T_EX and `language.dat.lua` used by LuaT_EX.

All the three are placed in

```
$TEXMFSYSVAR/tex/generic/config,
```

for example in

```
/usr/local/texlive/2017/texmf-var/tex/generic/config,
```

while a static copy also exists in

```
$TEXMFDIST/tex/generic/config,
```

but that one contains all languages and only makes sense if you have a complete checkout from SVN repository for example.

```
% from hyphen-norwegian:
bokmal loadhyph-nb.tex
=norwegian
=norsk
nynorsk loadhyph-nn.tex
```

Figure 2 language.dat

```
% from hyphen-norwegian:
\addlanguage{bokmal}{loadhyph-nb.tex}{2}{2}
\addlanguage{norwegian}{loadhyph-nb.tex}{2}{2}
\addlanguage{norsk}{loadhyph-nb.tex}{2}{2}
\addlanguage{nynorsk}{loadhyph-nn.tex}{2}{2}
```

Figure 3 language.def

```
-- from hyphen-norwegian:
['bokmal'] = {
    loader = 'loadhyph-nb.tex',
    lefthyphenmin = 2,
    righthyphenmin = 2,
    synonyms = { 'norwegian', 'norsk' },
    patterns = 'hyph-nb.pat.txt',
    hyphenation = 'hyph-nb.hyp.txt',
},
['nynorsk'] = {
    loader = 'loadhyph-nn.tex',
    lefthyphenmin = 2,
    righthyphenmin = 2,
    synonyms = { },
    patterns = 'hyph-nn.pat.txt',
    hyphenation = 'hyph-nn.hyp.txt',
},
```

Figure 4 language.dat.lua (the keyword loader is not used anywhere at the moment)

The files `language.dat` and `language.def` are used both by 8-bit and native Unicode engines. On one hand this simplifies things, while on the other it makes very little sense to include sanskrit patterns in pdfTeX – the package doesn't load any patterns in that case any way. It might be that there will be multiple copies of `language.dat` for different engines in future.

1.2 MiKTeX

In MiKTeX application **MiKTeX Options** which you can run from Start menu (*Maintenance* → *Settings* or *Maintenance (Admin)* → *Settings (Admin)*) or from command line (`mo.exe` or `mo_admin.exe`) you can select which languages you want to include into TeX formats.

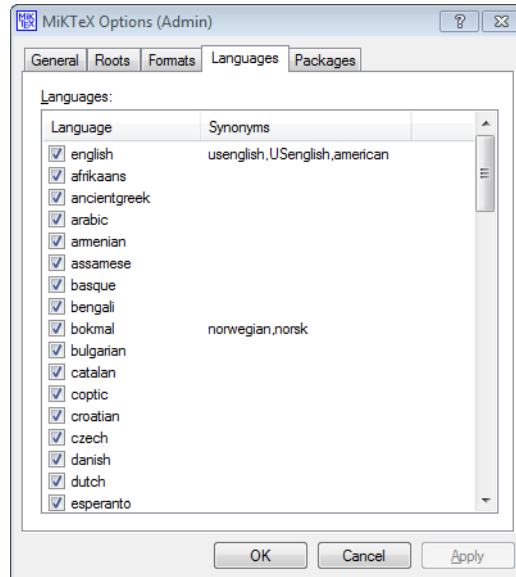


Figure 5 user interface to select the desired languages

Based on that selection MiKTeX creates three files `language.dat`, `language.def` and `language.dat.lua` in directory

```
C:\ProgramData\MiKTeX\2.9\tex\generic\config
```

The contents of those files are the same as described in TeX Live section, but the needed information about names of languages and required files comes from

```
C:\Program Files\MiKTeX 2.9\MiKTeX\config\languages.ini
```

instead of special hyphen-`<foo>` packages.

(Exact location might depend on where you install MiKTeX and whether you install it for all users or just for yourself. It may also depend on Windows version. The information above holds for MiKTeX 2.9 on Windows 7 with system-wide installation.)

```
[bokmal]
loader=loadhyph-nb.tex
lefthyphenmin=2
righthyphenmin=2
synonyms=norwegian,norsk
patterns=hyph-nb.pat.txt
```

```
[nynorsk]
loader=loadhyph-nn.tex
lefthyphenmin=2
righthyphenmin=2
patterns=hyph-nn.pat.txt
```

Figure 6 languages.ini

List of languages in TeX Live 2017:

collection-langarabic
 hyphen-arabic
 arabic
 hyphen-farsi
 farsi,persian
collection-langchinese
 hyphen-chinese
 pinyin
collection-langcyrillic
 hyphen-belarusian
 belarusian
 hyphen-bulgarian
 bulgarian
 hyphen-churchslavonic
 churchslavonic
 hyphen-mongolian
 mongolian
 mongolianlmc
 hyphen-russian
 russian
 hyphen-serbian
 serbian
 serbianc
 hyphen-ukrainian
 ukrainian
collection-langczechslovak
 hyphen-czech
 czech
 hyphen-slovak
 slovak
collection-langenglish
 hyphen-english
 ukenglish,british,UKenglish
 usenglishmax
collection-langeuropean
 hyphen-armenian
 armenian
 hyphen-croatian
 croatian
 hyphen-danish
 danish
 hyphen-dutch
 dutch
 hyphen-estonian
 estonian
 hyphen-finnish
 finnish
 hyphen-friulan
 friulan
 hyphen-hungarian
 hungarian
 hyphen-icelandic
 icelandic
 hyphen-irish
 irish
 hyphen-kurmanji
 kurmanji
 hyphen-latin
 classicallatin
 latin
 liturgicallatin
 hyphen-latvian
 latvian
 hyphen-lithuanian
 lithuanian
 hyphen-norwegian
 bokmal,norwegian,norsk
 nynorsk
 hyphen-occitan
 occitan
 hyphen-piedmontese
 piedmontese
 hyphen-romanian
 romanian
 hyphen-romansh
 romansh
 hyphen-slovenian
 slovenian,slovene
 hyphen-swedish
 swedish
 hyphen-turkish
 turkish
 hyphen-uppersorbian
 uppersorbian
 hyphen-welsh
 welsh
collection-langfrench
 hyphen-basque
 basque
 hyphen-french
 french,patois,français
collection-langgerman
 hyphen-german
 german
 ngerman
 swissgerman
collection-langgreek
 hyphen-greek

monogreek
greek,polygreek
hyphen-ancientgreek
ancientgreek
ibycus
collection-langitalian
hyphen-italian
italian
collection-langother
hyphen-afrikaans
afrikaans
hyphen-coptic
coptic
hyphen-esperanto
esperanto
hyphen-ethiopic
ethiopic,amharic,geez
hyphen-georgian
georgian
hyphen-indic
assamese
bengali
gujarati
hindi
kannada
malayalam
marathi

oriya
panjabi
tamil
telugu
hyphen-indonesian
indonesian
hyphen-interlingua
interlingua
hyphen-sanskrit
sanskrit
hyphen-thai
thai
hyphen-turkmen
turkmen
collection-langpolish
hyphen-polish
polish
collection-langportuguese
hyphen-portuguese
portuguese,portuges
collection-langspanish
hyphen-spanish
spanish,espanol
hyphen-catalan
catalan
hyphen-galician
galician