

Steps for transliteration without using ems

- setup srilm
 - download srilm
 - execute: make world
- setup mgiza
 - copy binaries to mosesdecoder/tools
- run training for transliteration

Execute command to train transliteration:

```
../mosesdecoder/scripts/Transliteration/train-transliteration-module.pl --corpus-f <foreign text>  
--corpus-e <target text> --alignment <path to aligned text> --moses-src-dir <moses decoder  
path> --external-bin-dir <external tools> --input-extension <input extension>--output-extension  
<output-extension> --srilm-dir <srilm binary path> --out-dir <path to generate output files>
```

for eg: en-hi

```
../mosesdecoder/scripts/Transliteration/train-transliteration-module.pl --corpus-f  
~/moses-home/corpus/train.clean.en --corpus-e ~/moses-home/corpus/train.clean.hi --alignment  
~/moses-home/working/train/model/aligned.grow-diag-final-and --moses-src-dir  
~/moses-home/mosesdecoder --external-bin-dir ~/moses-home/mosesdecoder/tools  
--input-extension en --output-extension hi --srilm-dir ~/moses-home/srilm/bin/i686-m64/ --out-dir  
~/moses-home/translit-en-hi
```

- Train moses with transliteration option on

```
nohup nice train-model.perl -root-dir train -corpus <path to parallel corpus> -f <foreign> -e  
<target> -alignment grow-diag-final-and -reordering msd-bidirectional-fe -lm 0:3:<path to lm>:8  
-external-bin-dir <external tools> -post-decoding-translit yes -transliteration-phrase-table <path  
to transliteration phrase table> >& training.out &
```

for eg: for language pair en-hi

```
nohup nice ~/moses-home/mosesdecoder/scripts/training/train-model.perl -root-dir train -corpus  
~/moses-home/corpus/train.clean -f en -e hi -alignment grow-diag-final-and -reordering  
msd-bidirectional-fe -lm 0:3:~/moses-home/lm/mono.ken.blm.hi:8 -external-bin-dir  
~/moses-home/mosesdecoder/tools -post-decoding-translit yes -transliteration-phrase-table  
~/moses-home/translit-en-hi/model/phrase-table.gz >& training.out &
```

- First pass decoding to generate output oov file and translation output without transliteration

```
nohup nice <path to moses> -f <moses.ini file> -output-unknowns <path to oov file to be output>  
< <path to test input> > <path to test output> 2> <path to trace output>
```

for eg:

```
nohup nice ~/moses-home/mosesdecoder/bin/moses -f filtered-testset/moses.ini  
-output-unknowns oov.out< test.head.en > translated.head.hi 2> translated.head.out
```

- Second pass decoding to transliterate to the output

```
./post-decoding-transliteration.pl --moses-src-dir <moses decoder> --external-bin-dir <external  
tools> --transliteration-model-dir <transliteration model> --oov-file <oov file obtained in previous  
step> --input-file <translated file obtained in previous step> --output-file <output translated file>  
--input-extension en --output-extension hi --language-model <path to language model>  
--decoder <moses executable>
```

for eg:

```
./post-decoding-transliteration.pl --moses-src-dir "~/moses-home/mosesdecoder"  
--external-bin-dir ~/moses-home/tools --transliteration-model-dir ~/moses-home/translit-en-hi  
--oov-file ~/moses-home/working-translit/oov_input --input-file  
~/moses-home/working-translit/translated.head.hi --output-file  
~/moses-home/working-translit/translated.head.out.hi --input-extension en --output-extension hi  
--language-model ~/moses-home/lm/mono.ken.blm.hi --decoder  
~/moses-home/mosesdecoder/bin/moses
```