

Poorman's Hangul Jamo Input Method

pmhanguljamo.sty

Kangsoo Kim

30 Jan 2020 version 0.3.1

Contents

1	Introduction	1
2	Usage	2
2.1	Loading the package	2
2.2	Commands and Environment Provided	2
2.3	Preamble of Document	3
3	Transliteration Rule	4
3.1	Tone Marks and Syllable Serapator	4
3.2	Consonants	4
3.3	Vowels	5
3.4	Compatibility Jamos	5
4	Proper Fonts	6
5	Examples	7
5.1	Modern Hangul	7
5.2	pre-1933 Hangul	7
6	The RRK input method	8
6.1	Transliteration Rule of RRK	8
6.2	Examples of RRK method	9
7	Further Information	10
8	Acknowledgement	10

1 Introduction

This \LaTeX package provides Hangul¹ *transliteration* input method, which allows to typeset Korean Letters (Hangul) with the help of proper fonts. The name comes from “Poorman’s Hangul Jamo Input Method.” It is mainly for the people who have a system without Korean IM, but want to typeset Hangul in their document. Not only modern

¹Hangul is the Korean alphabet to write the Korean language. In both South and North Korea, the standard writing system uses Hangul.

Hangul, but so-called “Old Hangul” characters that uses the lost letters such as ‘Arae-A’, ‘Yet Jeung’ or ‘Pan-Sios’ etc. can also be typeset.

X_YLaTeX or LuaLaTeX is required (X_YLaTeX recommended). The legacy pdfTeX is not supported. The Korean Language supporting packages such as xetexko or luatexko (in the ko.T_EX bundle) or polyglossia under X_YLaTeX are recommended, but without them typesetting Hangul is of no problem with this package pmhanguljamo.

2 Usage

2.1 Loading the package

Put the following line in your preamble:

```
\usepackage{pmhanguljamo}
```

The option [RRK] or [rrk] can be fed, in which case the *Standard Korean Romanization Transliteration* input method is activated. In section 6, we will explain about it. Without option, the default input method of this package is used, that is explained in sections 3-5. They are incompatible with each other.

2.2 Commands and Environment Provided

The main command to typeset Hangul is \jamoword.

Input: \jamoword{myec dan/e/ga or su iss/svb/ni/da/.}
Result: 몇 단어가 올 수 있습니다.

The argument of the command consists of roman alphabets and a few marks that represent the transliteration of Hangul Jamo. The transliteration rule will be explained in the section 3 (page 4).

And an environment named *jamotext* is provided. In the environment, multiple paragraphs can be located. So, it is suitable to typeset longer text stuffs. For example,

```
\begin{jamotext}  
jug/nvn nar/gga/ji ha/nvr/vr u/re/re \\  
han jem bu/ggv/rem ebs/gi/rvr  
  
ip/sai/ei i/nvn ba/ram/ei/do \\  
na/nvn goi/ro/ue/haiss/da/.  
\end{jamotext}
```

The result will be like this:

죽는 날까지 하늘을 우러러
한 점 부끄럼 없기를
앞새에 이는 바람에도
나는 괴로워했다.

2.3 Preamble of Document

Without any other packages, you can type Hangul. Note that proper fonts should be declared for Hangul, using fontspec.

```
\documentclass{article}
\usepackage{fontspec}
\usepackage{pmhanguljamo}

\begin{document}
\fontspec[Script=Hangul]{UnBatang.ttf}
\jamoword{an/nyex/ha/sei/yo}
\end{document}
```

UnBatang.ttf is a font shipped in unfonts-core package of T_EX Live. Not all Korean fonts can be available to typeset so-called ‘Old Hangul’, because they happen to lack the feature of composing Korean Syllables from conjoining Jamo codes. Fortunately, we have a few more fonts besides UnBatang that are *proper* and freely available, which are listed in the section 4. In T_EX Live distribution, UnBatang is the only *proper* one.

The polyglossia package provides Korean language support. This package pmhanguljamo can be used with it. The following is an example.

```
\documentclass{article}
\usepackage{polyglossia}
\setmainlanguage{english}
\setotherlanguage{korean}
\newfontfamily\hangulfont{UnBatang.ttf}[Script=Hangul]
\usepackage{pmhanguljamo}

\begin{document}
\begin{korean}
\jamoword{han/gvr/vr si/hem/hab/ni/da/.}
\end{korean}
\end{document}
```

Or the ko.T_EX package bundle can be used together.

```
\documentclass{article}
\usepackage{kotex}
\usepackage{pmhanguljamo}
\setmainhangulfont{UnBatang.ttf}[Script=Hangul]

\begin{document}
\jamoword{na/ras/mar/ss@/mi}
\end{document}
```

In any case, the option [Script=Hangul] should be given to the font settings for Hangul. For more information, see section 4.

The Korean writing system does not have hyphenation rules. The line breaking can occur just after almost every syllable. Therefore, without Korean language supporting packages the line breaking and justification of paragraphs will not be satisfactory.

However when just single or a few Hangul characters are needed, for example in the case of typesetting the author’s name in Hangul, this package provides a casual way to print them out.

3 Transliteration Rule

3.1 Tone Marks and Syllable Serapator

This package assumes that every Hangul syllable has Tone Mark which must be explicitly put at the end of it. Table 1 shows the input scheme of them.

Table 1: Tone Marks	
input	description
/	for <i>PyeongSeong</i> , no dot
;	for <i>GeoSeong</i> , single dot on the left of a character, [U+302E]
:	for <i>SangSeong</i> , double dots on the left of a character, [U+302F]

The tone marks were used in the 15th century, but in the modern Korean they are out of use. Nevertheless the / mark in this package is also used in order to depict the completion of composing a syllable — i.e., syllable separator. So DO NOT miss it.

However, it is possible to omit ‘/’ mark at the end of a word, i.e., just before <space>, or at the end of the argument of \jamoword. Note that before the punctuation mark — period , comma, etc., omitting / is not permitted.

3.2 Consonants

Table 2 shows the transliteration rule for the consonants which are used in the leading and trailing position of a syllable.

Table 2: Consonants							
ㄱ	g	ㄴ	n	ㄷ	d	ㄹ	r
ㅁ	m	ㅂ	b	ㅅ	s	ㅇ	x
ㅈ	j	ㅊ	c	ㅋ	k	ㅌ	t
ㅍ	p	ㅎ	h	ㅇ	q	ㅍ	z
ㅇ	f		w				

1. The corresponding letters “g, n, d, m, b, s, j, k, t, p, h” are the same as Korean Romanization Method (RRK).²
2. The letter for ‘ㄹ’ is allocated to r not l. The latter is reserved for another purpose, cf. item 7.
3. The letter ‘c’ is for ‘ㅊ’. In the RRK, it is ‘ch’ there.
4. Note that the letter for ‘ㅇ’ is x, which is a bit strange, but commonly accepted ‘ng’ was not to be chosen. It should be reserved for ‘ㅇ’.

²The *Revised Romanization of Korean* 1995:2000 is the official Korean language romanization system in South Korea.

5. And the letters ‘q, z, f’ are more freely selected. These letters are the lost ones in Modern Korean.
6. The letter ‘w’ is chosen to present *Choseong Filler*.
7. The letters ‘l’ and ‘lr’ are prepared for very rare case. In the 15th century Korean, ‘ㅌ’ and ‘ㅎ’ were distinguishable. ‘sl’ is chosen for the SIOS with left branch being longer, and ‘slr’ means right branch longer. The same is the case ‘ssl’, ‘sslr’, ‘jl’, ‘jlr’ etc.
8. The ‘light labial consonants’, ㅂ, ㅍ in the Old Hangul can be input as ‘bx, px, bbx, mx’.
9. In the Modern Korean, only ‘gg, dd, bb, ss, jj’ are additionally permitted in the leading consonant position, But in the Old Hangul this limitation was not applied. So *Choseong* can have multiple consonants, e.g., ‘bsg’, ‘bsd’, ‘ss’, ‘sd’ etc.
10. The leading ‘zero consonant’ can be omitted. So if a syllable begins with vowel, then the leading ‘o’ will appear. You can write the leading ‘x’ of a syllable or not. For example, ‘\jamoword{xax}’ and ‘\jamoword{ax}’ will generate the same result ‘ㅇ’.

3.3 Vowels

Table 3 shows the transliteration rule for vowels.

Table 3: Vowels					
ㅏ a	ㅓ e	ㅗ o	ㅜ u	ㅡ v	
ㅣ i	ㅛ @				
ㅑ ya	ㅕ ye	ㅙ yo	ㅠ yu	ㅚ @@	
ㅓ ay, ai	ㅓ ey, ei	ㅓ oy, oi	ㅓ uy, ui	ㅓ vy, vi	
ㅓ yay, yai	ㅓ yey, yei	ㅓ yoi	ㅓ yui	ㅓ @i	
ㅓ oa	ㅓ oay, oai	ㅓ ue	ㅓ uey, uei		

1. The rule for vowels may look unfamiliar. But it is simple and easily rememberable.
2. The letter ‘v’ is never considered to be a vowel. But this package determined it to stand for the vowel ‘ㅡ’. In fact, the roman alphabets fall short of the vowels. This allocation is arbitrary.
3. The lost ‘Arae-A’ is allocated to ‘@’, and the rare ‘Double Arae-A’ to ‘@@’, which are forgotten vowels in Modern Korean.

3.4 Compatibility Jamos

The compatibility Jamo block in Unicode contains several Hangul Jamo characters, which cannot be used to compose Hangul syllables, and are considered independent Jamo characters respectively.

To input the compatibility Jamos in the arguments of \jamoword or jamotext, you are to type them in their capital letters (upper cases). And every compatibility Jamo has

to be separated with syllable separator — /. Because the @ character has no upper case, write W instead of @ to display compatibility ‘Arae-A’.

Here is an example:

```
\jamoword{W/nvn a/rai/a/ra/go bu/rv/go A/oa bi/svs/ha/gei irg/
nvn/da}
、는 아래아라고 부르고 ㅏ와 비슷하게 읽는다
```

4 Proper Fonts

In order to typeset Hangul characters successfully with the input method of this package, the help of the *proper* font is crucial. What we call *proper* font is the font that has the Opentype feature of ‘Script Hangul’, that is, the feature of conjoining jamo characters. A Truetype font may have this Opentype feature. Otherwise, it is regarded as an *improper* font.

Not so many fonts are *proper*. The fonts or font families listed in the following lines are what we know *proper*. They can be downloaded from the url and used freely. And note that UnBatang is already installed in your T_EX Live installation.

Hamchorom LVT Hamchorom fonts by Hancom Inc., and their variants with GSUB tables by KTS (the Korean T_EX Society). <https://github.com/dohyunkim/hcr-lvt>

KoPub World free fonts by Korea Publisher Society. <http://www.kopus.org/biz/electronic/font.aspx>

Malgun Gothic One of Windows 10 system font.

Nanum Yet Hangul Two font families from Naver. Nanum Myeongjo Yet Hangeul and Nanum Barun Gothic Yet Hangeul. <https://hangeul.naver.com/2014/archaicword>

Noto CJK Google’s Noto CJK fonts. Or Adobe’s Source Han fonts. <https://www.google.com/get/noto/help/cjk/>

Un Batang UnBatang.ttf. on CTAN. <https://ctan.org/pkg/unfonts-core>

To use one of these fonts, do not forget to provide the [Script=Hangul] option.³ For example, you can specify Hamchorom Batang LVT font as:

```
\newfontfamily\pmjamofont{HCR Batang LVT}[Script=Hangul]
```

and then issue \pmjamofont command in the scope of Hangul to be typeset. It is recommended that the *proper* font should be used consistently.

In case the font you take is not *proper*, you are to meet the warning message at the console and log file,

```
Package fontspec Warning: (fontspec)
```

```
Font "batang" does not contain requested Script "Hangul".
```

However, when X_YL_AT_EX is concerned, it is known that the *Modern* Korean Characters (*Hangul Precomposed Syllables* in the Unicode block [U+AC00] – [U+D7A3]) can be rendered with this kind of font, in spite of the warning message.

³If the document is processed with Lua_LA_TE_X, [Script=Hangul,Renderer=Harfbuzz] is recommended. The Lua_LA_TE_X version should be over 1.11, that is, the L_AT_EX format of LuaHB_TE_X is needed.

5 Examples

Now, we will show a few practical examples, where the *korean* environment of polyglossia package is used. The preamble is like this:

```
\usepackage{pmhanguljamo}  
\usepackage{polyglossia}  
\setotherlanguage{korean}  
\newfontfamily\hangulfont{UnBatang.ttf}[Script=Hangul]
```

5.1 Modern Hangul

```
\begin{verse}  
\begin{korean}  
\begin{jamotext}  
na bo/gi/ga yeg/gye/ue \\  
ga/sir ddai/ei/nvn \\  
mar ebs/i go/i bo/nai dv/ri/u/ri/da/.  
  
yex/byen/ei yag/san \\  
jin/dar/rai ggoc \\  
a/rvm dda/da ga/sir gir/ei bbu/ri/u/ri/da/.  
  
ga/si/nvn ger/vm ger/vm \\  
noh/in gv ggoc/vr \\  
sa/bbun/hi jv/rye/barb/go ga/si/ob/so/se  
  
na bo/gi/ga yeg/gye/ue \\  
ga/sir ddai/ei/nvn \\  
jug/e/do a/ni nun/mur hvr/ri/u/ri/da/.  
\end{jamotext}  
\end{korean}  
\end{verse}
```

나 보기가 역겨워
가실 때에는
말 없이 고이 보내 드리우리다.
영변에 약산
진달래 꽃
아름 따다 가실 길에 뿌리우리다.
가시는 길음 걸음
놓인 그 꽃을
사뿐히 즈려밟고 가시옵소서
나 보기가 역겨워
가실 때에는
죽어도 아니 눈물 흘리우리다.

5.2 pre-1933 Hangul

A 15th century text with tone marks (*BangJeom*).

```

\begin{korean}
\begin{jamotext}
na/ras;mar:ss@/mi; dyuq/guig;ei; dar/a;
mun/jj@x;oa;ro; se/rv s@m@s/di; a/ni;h@r/ss@i;
i;ren jyen/c@;ro; e/rin; b@ig;syeq;i;
ni/rv/go;jye; horf; bai;
i/sye;do; m@c@m;nai: jey bdv;dvr; si/re; pye/di;
mod:h@rf no;mi; ha/ni;ra;.
\end{jamotext}
\end{korean}

```

나랏말싸미 등꺽에 달아 문중와로 서르스 못디 아니홀씨 이런 전초로
어린 북성이 니르고져 홀 배 이셔도 마춤내 제 ㅁ들 시러 퍼디 묻흠
노끼 하나라.

And another example of 17th century, from 두시언해, *Selected Translations of Du Fu's Poems*, without tone marks.

```

\begin{verse}
\begin{korean}
\begin{jamotext}
g@/r@/mi p@/r@/ni sai de/ug h@i/o/,\\
moy/hi pe/re/h@/ni gos/ bi/ci bvr bvd/n@n d@s/do/da/.\\
ors bo/mi bon/d@in sdo di/na/ga/n@/ni \\
e/nv na/ri i do/ra/gar h@i/o/.
\end{jamotext}
\end{korean}
\end{verse}

```

구름미 푸르니 새 더욱 희오,
되히 퍼러흐니 곳 비치 불 뜰는 듯도다.
웁 보미 본딘 또 디나가느니
어느 나리 이 도라갈 희오.

6 The RRK input method

In chapter 3 section 8 of *Revised Romanization of Korean* (2000), the ‘transliteration method’ of Hangul romanization is declared. It is exceptionally permitted when it is required to restore the original Hangul writings, while the phonetic transcription method is standard. So we can make use of it as an input method.

If the package option [RRK] or [rrk] is given, the RRK method is activated. In this case, only ‘Modern Hangul’ can be available.

6.1 Transliteration Rule of RRK

Table 4 and 5 show the RRK transliteration rule.

1. When the leading consonant ‘ㅇ’ is placed at the head of a word, it can be omitted. Otherwise the hyphen character (-) should be put. For example,
`\jamoword{annyeonghase-yo}` 안녕하세요
`\jamoword{jung-ang jungang}` 중앙준강
`\jamoword{ha-yeossseubnida}` 하였습니다

Table 4: RRK Consonants

ㄱ	g	ㄲ	kk	ㅋ	k	ㄴ	d	ㄷ	tt	ㅌ	t
ㅂ	b	ㅃ	pp	ㅍ	p	ㅈ	j	ㅊ	jj	ㅊ	ch
ㅅ	s	ㅆ	ss	ㅎ	h						
ㄴ	n	ㅁ	m	ㅇ	ng	ㄹ	l				

Table 5: RRK Vowels

ㅏ	a	ㅑ	eo	ㅓ	o	ㅕ	u	ㅡ	eu	ㅣ	i
ㅗ	ae	ㅓ	e	ㅜ	oe	ㅠ	wi				
ㅗ	ya	ㅑ	yeo	ㅜ	yo	ㅠ	yu	ㅑ	yae	ㅑ	ye
ㅘ	wa	ㅑ	wae	ㅜ	wo	ㅠ	we	ㅑ	ui		

- The hyphen character - can be used to forcibly separate syllables. And the special character / is taken as another syllable separator. They are exchangeable.
`\jamoword{haeb/ssal haeb-ssal}` 햅쌀 햅쌀
- The six punctuation marks, . , ! ? ; : can be located in the Hangul scope. The others are not allowed.
`\jamoword{a! malg-eun haneul, si-wonhan baram.}`
아! 맑은 하늘, 시원한 바람.
The \LaTeX quotes, ``', ' and ' cannot be used within Hangul input scope. They should be located outside. However you can use unicode quotes, " and ", which can be input with syllable separator. For example,
`\jamoword{"/hangeul"//ilan hangug-eoleul jeogneun geul-ida.}`
“한글”이란 한국어를 적는 글이다.
In this case, the syllable separator / cannot be replaced by hyphen character.
- The leading consonant ‘ㄹ’ is written with either l or r, while the trailing one should be l.
`\jamoword{noraereul nolaeleul}` 노래를 노래를
- We made it possible to use ‘Arae-A’ with the RRK method. Though it is not present in Modern Hangul writing system, typesetting ‘Jeju language’ or some pre-1933 Korean texts will be made easy with this. The ‘Arae-A’ is to be input as ‘@’ character, and the ‘Double Arae-A’ ‘@@’. There is one exception. The one word with one letter that has only leading ‘@’, i.e., ㄹ should be input by ‘x@’. If the syllable has trailing consonants or is part of a word, just @ is good.
`\jamoword{h@ngeul @@nam-eun}` 한글 ㄹ 남은.

6.2 Examples of RRK method

The text is one of the most famous poem in Korean, 진달래꽃 *Azalea* by Kim So-wol.⁴ In section 5.1, we showed default way of inputting the poem. The *korean* environment and font settings are the same as in section 5.

```
\begin{verse}
\begin{korean}
```

⁴An English translation of this poem can be read in Wikipedia, https://en.wikipedia.org/wiki/Kim_Sowol.

```

\begin{jamotext}
na bogiga yeoggyeo-wo \\
gasil ttae-eneun \\
mal eobs-i go-i bonae deuli-ulida.

yeongbyeon-e yagsan \\
jindallae kkoch \\
aleum ttada gasil gil-e ppuli-ulida.

gagineun geol-eum geol-eum \\
noh-in geu kkoch-eul \\
sappunhi jeulyeobalbgo gasi-obseseo

na bogiga yeoggyeo-wo \\
gasil ttae-eneun \\
jug-eodo ani nunmul heulli-ulida.
\end{jamotext}
\end{korean}
\end{verse}

```

나 보기가 역겨워
 가실 때에는
 말 없이 고이 보내 드리우리다.
 영변에 약산
 진달래 꽃
 아름 따다 가실 길에 뿌리우리다.
 가시는 걸음 걸음
 놓인 그 꽃을
 사뿐히 즈려밟고 가시옵소서
 나 보기가 역겨워
 가실 때에는
 죽어도 아니 눈물 흘리우리다.

7 Further Information

For more examples and information about this package, please read the documentation [pmhanguljamo-kdoc.pdf](#) (in Korean).

8 Acknowledgement

Thanks go to yihoze at the KTUG board, there he commented, “The IMs I’ve tried such as Korean Windows Old Hangul IM or Saenaru IM were all somewhat inconvenient. ... In my opinion, for the foreign scholar who studies Korean language, especially Middle Age Korean language, this kind of transliteration method may be much more convenient and confident way of typesetting Korean than the IM solutions.”⁵

⁵http://www.ktug.org/xe/index.php?mid=KTUG_open_board&document_srl=240134