



WaitZar User's Guide

SCIM Wait Zar version 1.0.0

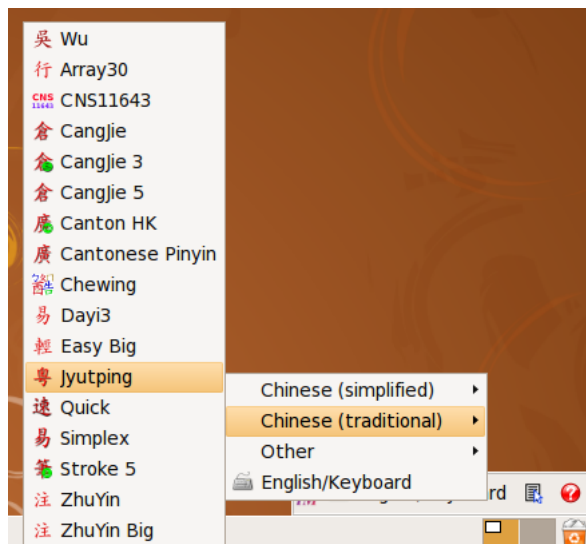
Table of Contents

1. Introduction to SCIM and WaitZar	2
2. Requirements & Installation	3
2.1. Installation Using Apt	3
2.2. Installing the .deb File Directly	4
2.3. Building & Installing from Source	4
2.4. Fonts & Configuration Extras	5
3. Basic Usage	8
3.1. Introduction to SCIM	8
3.2. Typing Your First Myanmar Sentence	8
3.3. Additional WaitZar Hotkeys	9
3.4. The Open Office Glitch	9
3.5. Changing the Output Encoding	10
3.6. A Note About config.txt	11
3.7. Adding custom words to mywords.txt	11
4. Finding a Word	12
4.1. The WaitZar Romanisation	12
4.2. Dictionary Lookup	19
5. Known Bugs	20
5.1. Pat-Sint Choices are Glitched	20
5.2. Alt+Tab Sometimes “Sticks”	20
5.3. Choices Stuck in the Lower-Left	20
5.4. The Console Looks Glitchy	21
A. License	22
B. How It Works	22

1. Introduction to SCIM and WaitZar

If you're a Linux user who knows how to write a complex script like Chinese or Thai, chances are you've used SCIM before. Even some European users prefer SCIM, although their languages do not require such a complex solution. The **scim-waitzar** project extends SCIM to support the Myanmar language.

Figure 1. A screenshot of SCIM running on Ubuntu, showing a wide variety of Chinese input method engines



WaitZar is a romanisation of Burmese designed specifically for fast typing on a standard keyboard. Given the fluid nature of the Myanmar unicode specification (i.e., the 4.0, 5.0, and 5.1 encodings are mutually incompatible) and the wide variety of *ad hoc* encodings (e.g., Zawgyi-One partial, Win Innwa's ASCII kludging, etc.), WaitZar provides a degree of solidity and flexibility above and beyond that provided by most other input schemes for Burmese.

The official project page of **scim-waitzar** is:

<http://scim-waitzar.googlecode.com/>

If you would like to keep casually up-to-date on WaitZar's Windows and Linux projects, you might consider joining our mailing list, which averages about one announcement per month:

<http://groups.google.com/group/waitzar>

2. Requirements & Installation

WaitZar will run on almost any Linux distribution that also supports SCIM. In fact, by installing the **scim** package (or building SCIM from source), you should already have everything you need. WaitZar is recommended for Ubuntu Linux, but it has also been tested on Debian Linux. Please inform us if **scim-waitzar** works (or breaks!) on your favorite brand of Linux.

Unlike the Windows version, WaitZar under Linux requires a significant amount of installation effort. The most stable method is to build **scim-waitzar** from source; see Section 2.3. However, a great many build steps are simplified if your system supports Debian packages. See Section 2.1 if you're used to using apt-get, and Section 2.2 if you'd prefer to (manually) use dpkg. Finally, Section 2.4 explains several very important things to consider after installation (using any of the three methods.) Until WaitZar gets fully integrated into Ubuntu's *Language* manager, you will have to consider these.

You should have at least one of the following Unicode 5.1-enabled fonts installed on your system. Please see Section 2.4 for more information on how to install fonts.

- Padauk
- Parabaik
- Myanmar 3

You should install one of these fonts (for example, **ttf-sil-padauk**) in order to view this document properly.



Note Regarding Fonts

Traditionally, Burmese fonts have defined their own *ad hoc* encodings, which has caused a lot of confusion among novice computer users and experts alike. Even previous versions of the Unicode encoding (for example, 4.0 and 5.0) had non-negligible errors and incompatibilities. Fortunately, Unicode 5.1 has resolved most of these, and the three fonts listed all render this document with no errors.

Developers, please note that, when discussing character encoding incompatibilities (in, say, the "WaitZar User Interface Specification"), the documentation will use the Zawgyi-One font & encoding, due to its unambiguous representation of variant forms.

2.1. Installation Using Apt

The recommended way of installing WaitZar is to use **apt-get** or the Synaptic Package Manager. In order to do this, you will have to add a line to the sources file. To do this, open a console and type:

```
gksudo gedit /etc/apt/sources.lst
```

...and add the following repository lines at the very end. Note that these links are temporary. After we finish alpha testing, we'll put the **scim-waitzar** package at a more suitable location.

```
deb http://www.comp.nus.edu.sg/~sethhetu/scim_wz_1.0.0/ ./
deb-src http://www.comp.nus.edu.sg/~sethhetu/scim_wz_1.0.0/ ./
```

...then, save the file and close **gedit**, and type:

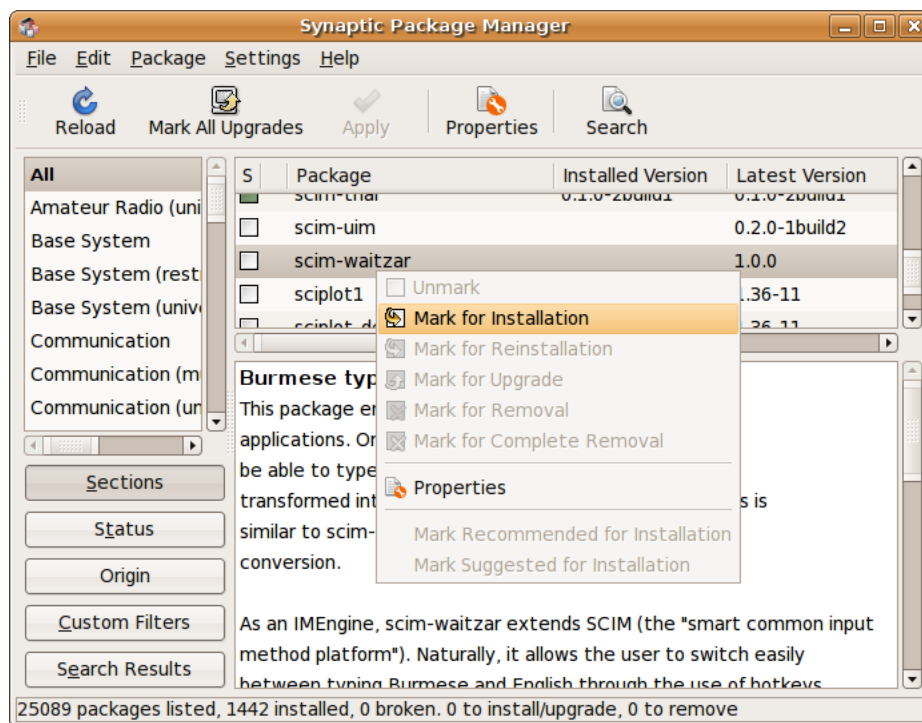
```
sudo apt-get update
```

This will update the list of packages. Finally, install **scim-waitzar**:

```
sudo apt-get install scim-waitzar
```

Of course, you can also use Synaptic at this point:

Figure 2. After reloading your repositories, you can simply install WaitZar like any other package.



2.2. Installing the .deb File Directly

In general, the apt repository will never go offline. Nevertheless, you can always download and install the .deb package manually. Grab the latest version from:

<http://code.google.com/p/scim-waitzar/downloads/list>

...and the latest version of libwaitzar from:

<http://code.google.com/p/waitzar/downloads/list>

...and then run the standard **dpkg** install command:

```
sudo dpkg -i libwaitzar_*.deb scim-waitzar_*.deb
```

...from the directory you downloaded the package to. You must then restart X (log out, log in). In case you are not familiar with **dpkg**, note that un-installation can be done from any directory, and requires only the logical name of the package:

```
sudo dpkg -r scim-waitzar
```

2.3. Building & Installing from Source

Building from source is actually quite simple, thanks to the GNU autotools. Here are the steps for the Ubuntu operating system, although they should work with any **automake**-compatible system. These instructions are for the "developer hat", since users will probably just use **apt-get** to install WaitZar. If you are wearing your "user hat", you should still be able to follow along.

1. Install necessary packages, preferably with **apt-get**:

```
sudo apt-get install subversion autoconf automake gettext libtool scim-dev g++ libwaitzar-dev
```

If apt cannot find libwaitzar-dev, you can build and install that from source as well, using instructions similar to those provided in this section.

2. Make a directory for the source files, and download them with **svn**:

```
mkdir ~/scim-waitzar
svn checkout http://scim-waitzar.googlecode.com/svn/trunk/ ~/scim-waitzar
cd ~/scim-waitzar
```

3. Reconfigure the system:

```
autoreconf
```

4. Run through the standard autotools lifecycle:

```
./configure
make
sudo make install
```

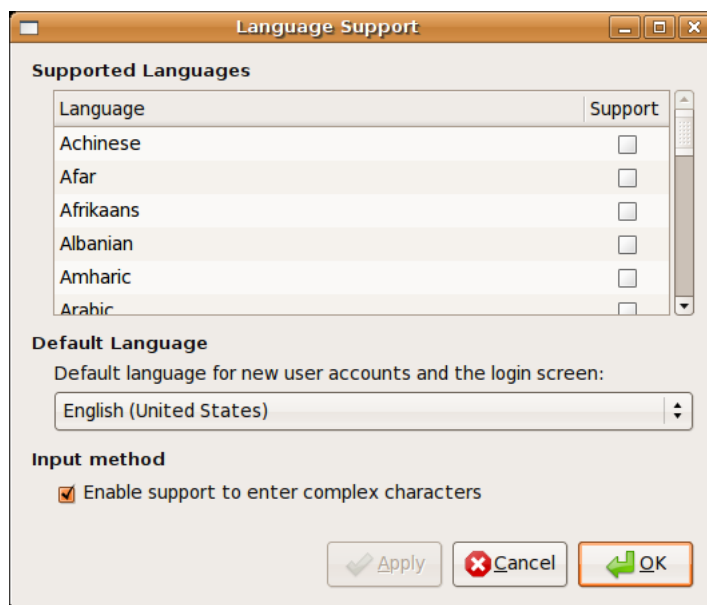
5. Log out, then log in again. WaitZar is now (almost) ready to use.

2.4. Fonts & Configuration Extras

Although most Linux programs have been properly internationalized, a great many users simply do not trust “tacking another layer” onto input (that layer being SCIM). Hence, simply installing WaitZar will not be enough—you’ll also have to enable it, play around with **fontconfig**, and probably install some fonts, too. Here are some helpful hints and tricks. It should go without saying that, although these techniques are harmless, we take no responsibility for any unintentional side effects they have on your system.

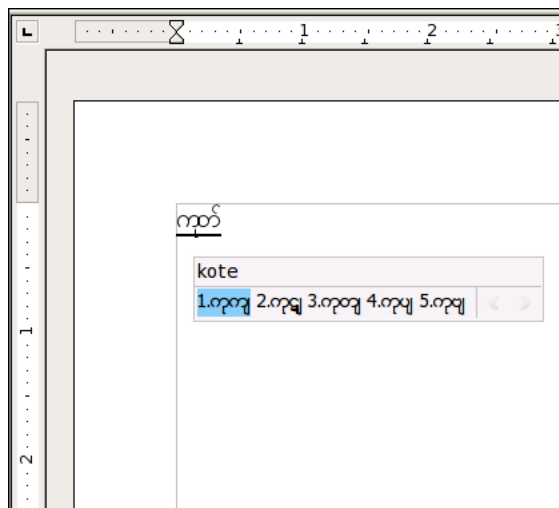
Enable Chinese — Making every application recognize SCIM is not easy; there are reports that “SCIM won't work in Open Office” which are only half true. Fortunately, there's a cheap trick that works well in Ubuntu: just enable Chinese, and the system will figure all this out for you. Click on “System” then “Administration”, then “Language Support”, and check the box marked “Chinese”. Several package files will be installed. Finally, check the box that reads “Enable support to enter complex characters”, and click “Ok”. You will probably have to restart. Note that, if you do not have at least one complex language installed (like Chinese), checking the box will have no effect. That is why we recommend enabling Chinese, specifically.

Figure 3. The “Language Support” window. Make sure the “complex characters” checkbox is checked.



Finagle FontConfig — SCIM does not dictate which font should be used for its main interface, relying on *fontconfig* like every other program does. That means the wrong font can be picked for the wrong encoding. For example, here is what WaitZar looks like if Zawgyi-One is chosen by mistake.

Figure 4. Trying to type “kote” with a mis-configured font selector. Note that this will not actually affect the typed word, just the list of choices.



We recommend listing at least one Unicode 5.1 font in *local.conf*’s “prefer” tags. We’ve created a sample file for you that also makes your text sharper, and has Chinese fonts listed properly. You can get it here:

<http://scim-waitzar.googlecode.com/svn/trunk/local.conf>

...copy it to

/etc/fonts/local.conf

...using, for example:

```
cd /etc/fonts && sudo cp local.conf local.bak  
sudo wget http://scim-waitzar.googlecode.com/svn/trunk/local.conf
```

You'll probably have to restart X. (Log out/Log in).

Install Some Fonts — You'll need at least one font installed to use WaitZar effectively. Some fonts have packages to make installation easier; for example, the *Padauk* font has **ttf-sil-padauk**. If no package exists, you have several options. The simplest is to make a user-level fonts directory:

```
mkdir ~/.fonts
```

...and then copy any fonts you download to that folder. You should not have to restart X, but you might need to restart any running application you want to use the font in.

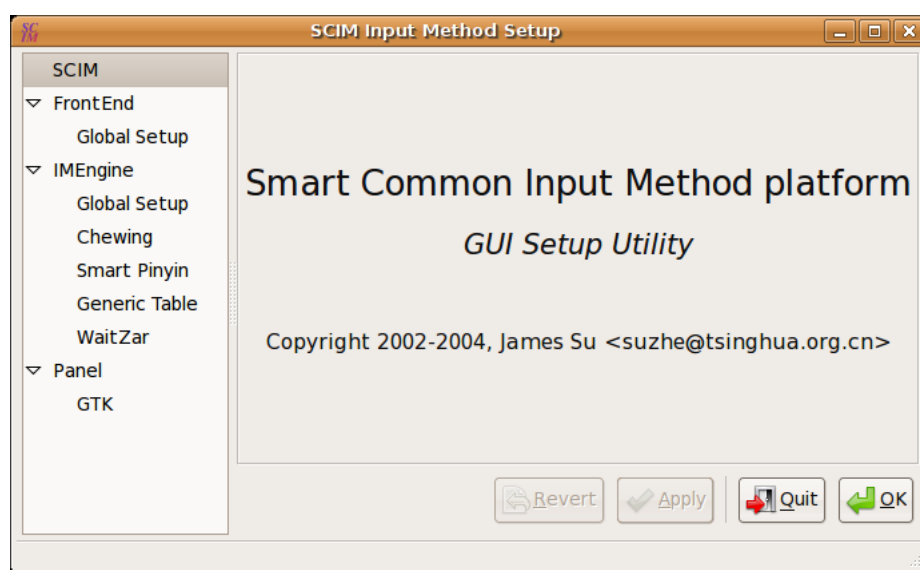
3. Basic Usage

3.1. Introduction to SCIM

After you install **scim-waitzar**, you should take some time to get used to SCIM's unique feel. You will see a small keyboard-shaped icon in your tray, right-click it, and choose “SCIM Setup”. This panel can be quite confusing for new users, so we'll briefly cover it here.

The “Front End → Global Setup” panel contains SCIM's hotkeys, and various global settings. You will probably want to change the hotkey from Ctrl+Space to *anything* else, if you are a programmer.

Figure 5. The SCIM setup window.



The next section (IMEngine) has settings for each “Input Method Engine”. Recalling that SCIM provides support for multiple languages, it makes sense that each language should have its own customizations. The final section, “Panel → GTK”, holds most of the visual settings for SCIM. I recommend changing the “Show” option (under “Toolbar”) to “Always”, at least until you get used to how SCIM behaves. Note that when you click “Ok” to exit the setup panel, SCIM will warn you that you might need to restart SCIM for any changes to take place. The sure-fire way to do this is to log out and then log in again. Optionally, you might right-click on the SCIM icon and choose “Reload Configuration”. In some cases, the operating system will reload SCIM for you. Regardless, you won't be playing around with the WaitZar settings that much anyway.

3.2. Typing Your First Myanmar Sentence

After configuring SCIM just the way you like it, start **gedit** and then look at the lower-right corner of your screen. You'll see the SCIM toolbar; move the mouse over it and it will expand to show that you are typing in English. Click on the language and choose “Burmese” to switch to WaitZar. (Note that if this panel is hidden, you can always click on the SCIM icon in your tray to switch to Burmese.) Then, type “ko”. You will see two possibilities presented to you.


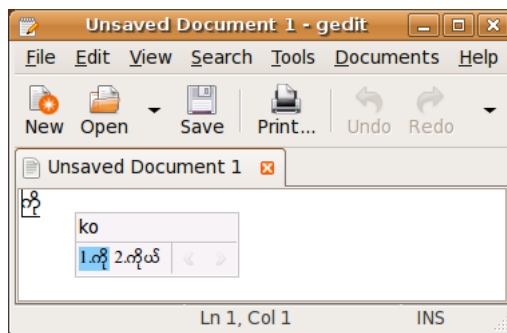
You will also see the word “” inside of **gedit**. This is one of the nice features of WaitZar on Linux: you can keep your eyes on the document you are typing, and only have to look at the list of choices when something doesn't match.

Figure 6. Choosing between two "ko"s.



Hit “space”, and the cursor will move right; the list of choices will disappear. Now, type “2”, and you will see the Burmese numeral “၂”. Finally, hit “Enter”, and the words you typed will no longer be underlined. Why did this happen? Basically, WaitZar “pre-edits” anything you type. Hitting left and right works correctly in WaitZar because it delays direct typing of words. Although this may take some getting used to, it is actually quite natural, and there are several shortcuts to save you time.

3.3. Additional WaitZar Hotkeys

By default, SCIM lets you switch to the *last* language you used by hitting Ctrl+Space. Hitting it again will switch to English.

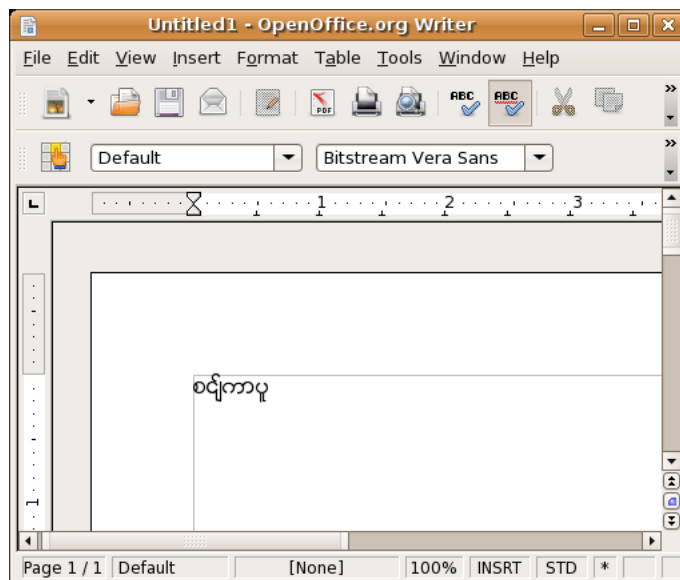
When you are typing in WaitZar, several other hotkeys also take effect:

- Pressing any letter will type that letter, and attempt to form a word.
- Pressing left and right will move the cursor. When you are typing a word, left and right switch between the various choices.
- Pressing “enter” will confirm the typed sentence. If you are typing a word, hitting “enter” is the same as hitting “space”.
- Pressing “space” will move the cursor right. If the cursor is all the way to the right, hitting “space” will confirm the typed sentence. When you are typing a word, hitting “space” will choose the currently-selected choice.
- Hitting the number keys (0-9) will type the burmese numeral (၀-၉). If you are typing a word, hitting a number will “shortcut” to that word. (Try typing “ko2” to see this work).
- Hitting the period or comma keys will type the Burmese full or half stop characters (“.” and “,”). If you have typed part of a sentence, this will also confirm the sentence.
- Hitting “esc” will cancel the sentence or word you are typing.
- Hitting “backspace” and “delete” will function as expected, deleting letters or words depending on the context.
- In the rare case that you are using an absolutely tiny display, hitting up and down will scroll the list of choices if there are too many to fit on the screen at once.

3.4. The Open Office Glitch

There are several glitches related to WaitZar; most are covered in Section 5. However, one particularly nasty bit occurs in Open Office Writer –to see it for yourself, change the font (in Open Office) to Padauk, type “singapore”, and then hit “space” twice to finalize it. The result is a mess of glyphs:

Figure 7. Open Office seems to like resetting the font for you, even when this is actually incorrect.



Fortunately, this is easy to fix: just select all the text and change the font to “Padauk” again. For some reason, Open Office reverts the font to *Bitstream Vera Sans*, an otherwise-excellent font that, unfortunately, cannot entirely handle Burmese just yet.

Even more fortunately, this only has to be done once: the first time you create a new document. Even if you delete everything on the page, Open Office can still manage to change the font back to Padauk the second time you enter Burmese text.

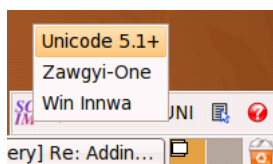
On a side note, this glitch is far worse on Windows, which doesn't provide a reasonable catch-all font. In this case, the first Burmese word you type in Open Office will actually be invisible!

3.5. Changing the Output Encoding

A great many websites use encodings that conflict with the established standard. Criticizing them will do no good; these *ad hoc* encodings were all developed to suit a particular niche, and they will likely be around until the end of time.

WaitZar supports two of these encodings: the Zawgyi-One font encoding, and the Win Innwa font family encoding. (Win Kalaw, for instance, is in the Win Innwa font family). To change the encoding, start typing in Burmese, and then click the “UNI” icon in the SCIM toolbar. Now choose the encoding you want. This will reset the next time you log in; if you want to change the encoding used by default, see Section 3.6.

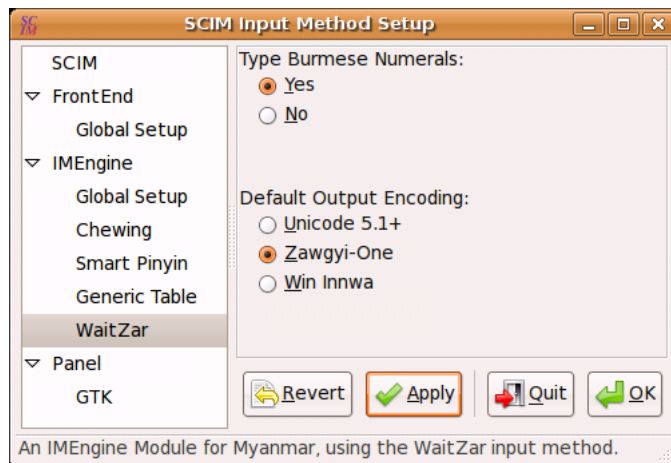
Figure 8. Changing the output encoding.



3.6. A Note About config.txt

For Windows users of WaitZar, it should come as a mildly unpleasant surprise that config.txt does nothing. The proper way to change WaitZar's configuration settings is through the SCIM settings panel, on the “WaitZar” tab. You can currently only change two settings: how Burmese numbers are typed, and the default output encoding. A screenshot of this panel follows.

Figure 9. WaitZar's settings: here, the user has chosen to type numbers in Burmese, and to start WaitZar with Zawgyi-One by default.



3.7. Adding custom words to mywords.txt

WaitZar's dictionary is drawn from online forums, books, and news articles, and it contains most of the words needed by most of the users. However, it is possible that you may find it lacking a word you need. Or, you might type a word so often (like your name) that you want to shorten it. For this, you will add lines to the file "mywords.txt" of the form:

```
myanmar = roman
```

For example:

```
မ = ma
```

Note that, for legacy reasons, you will have to enter the word you require in the Zawgyi-One encoding. If you have the word in a different encoding, the following web site provides excellent software for converting it:

<http://burglish.googlepages.com/fontconv.htm>

The only remaining detail is *which* mywords.txt file to edit. For global changes to WaitZar, you can change /usr/share/scim-waitzar/mywords.txt. For user-level changes, you should change ~/.scim-waitzar/mywords.txt. The latter way is preferred. You should use the UTF-8 encoding for this file; for example:

```
mkdir -p ~/.scim-waitzar
gedit --encoding UTF-8 ~/.scim-waitzar/mywords.txt
```

4. Finding a Word

Although WaitZar was designed to be natural to use, there will come a time when you do not know how to type a word in WaitZar. Fortunately, as of release 1.6 on Windows (and 1.0.0 on Linux), WaitZar now ships with version 2.0 of the WaitZar romanisation. This romanisation was designed to be simple and effective, and to make looking up unknown words easy. We will now discuss how this works.

4.1. The WaitZar Romanisation

The WaitZar romanisation was designed around three simple principles:

- There should be a few simple rules which most words follow.
- All words should be easy to guess, even if you don't know the rules.
- Each Burmese word gets exactly one English romanisation.

The first rule is covered in the next section (on the default romanisation), and the second rule is relevant throughout. The final rule may seem silly: why not allow တာ = “tar” or “thar” or “ttar”? Our main reason for enforcing this rule is efficiency. Once you learn that တာ = “thar” in WaitZar, then you will start to get annoyed if you see တာ when you type “tar” or “ttar”. Through decisions like this, we try to make our system agreeable to both novices and experts.

4.1.1. Default Romanisation

Most Burmese words can be thought of as an onset combined with a rhyme. For example: န, pronounced “non”, can be broken up into န = “n(a)” and ဝ = “on”. Words can actually be broken up further, into consonants, medials, vowels, tones, and finals, but a unique quality of onsets and rhymes is that they can be spelled the same way when they appear in different words. For example:

န → n		ဝ → on	
Myanmar	Roman	Myanmar	Roman
န	non	ဝ	non
န	nin	ဝ	khon
န	ne	ဝ	son
န	nay	ဝ	yon
န	nwal	ဝ	thon

The rhymes always follow their phonetic pronunciation; the onsets are more structured. The following rules apply to onsets:

1. The first sound of the Myanmar letter is the Roman letter.
2. If two onsets share the same first sound, the less common onset has an “h” after it. (For example: ခ is “d”, but ခ is “dh”)
3. Rarely-used letters won't add an “h”. (For example: ဂ and ဝ are both “n”, even though န is also “n”).

For reference, here is the current list of rhymes and onsets. Within a rhyme, a dash indicates where onsets may be plugged in. We apologize for the large size of these tables; we were forced to include them as images, since none of the Unicode-enabled fonts provides a convenient way for representing Burmese syllable fragments. The alternative was to require you to install the (non-compliant) Zawgyi-One font, which we felt was unreasonable.

Onsets							
<i>Myanmar</i>	<i>Roman</i>	<i>Myanmar</i>	<i>Roman</i>	<i>Myanmar</i>	<i>Roman</i>	<i>Myanmar</i>	<i>Roman</i>
က	k	တ	tt	ခ	dh	ဝ	w
ခ	kh	ဒ	dh	န	n	သ	th
ဂ	g	င	ht	ပ	p	ဟ	h
ဃ	gh	ဇ	ht	ဖ	ph	အ	
င	ng	ဇ	n	ဓ	v	ဥ	o
စ	s	ဗ	d	ဘ	b	ရှ	sh
ဆ	ss	ဏ	nh	မ	m	သျှ	sh
ဇ	z	တ	t	ယ	yh	လျှ	sh
ဈ	zh	ဆ	ht	ရ	y	ရွှ	shw
ည	ny	ဒ	d	လ	l		

Rhymes											
my	rom	my	rom	my	rom	my	rom	my	rom	my	rom
-	a	ၵ်း	ote	ၵ်း	ya	ၵ်း	ywae	ၵ်း	oi	ၵ်း	shae
ၵ်း/ၵ်း	ar	ၵ်း	ate	ၵ်း	yar	ၵ်း	yaw	ၵ်း	wet	ၵ်း	ywa
ၵ်း/ၵ်း	a	ၵ်း	at	ၵ်း	yi	ၵ်း	yan	ၵ်း	ut	ၵ်း	yaunote
ၵ်း/ၵ်း	arr	ၵ်း	at	ၵ်း	ye	ၵ်း	yant	ၵ်း	oot	ၵ်း	way
ၵ်း	i	ၵ်း	an	ၵ်း	yee	ၵ်း	yo	ၵ်း	on	ၵ်း	way
ၵ်း	e	ၵ်း	am	ၵ်း	yu	ၵ်း	yoe	ၵ်း	ont	ၵ်း	wal
ၵ်း	eet	ၵ်း	ai	ၵ်း	yuu	ၵ်း	yoe	ၵ်း	onn	ၵ်း	yon
ၵ်း	ee	ၵ်း	ohn	ၵ်း	yuu	ၵ်း	yone	ၵ်း	wot	ၵ်း	yot
ၵ်း/ၵ်း	u	ၵ်း	hite	ၵ်း	yay	ၵ်း	yote	ၵ်း	wan	ၵ်း	ywin
ၵ်း/ၵ်း	uu	ၵ်း	ine	ၵ်း	yay	ၵ်း	yone	ၵ်း	wann	ၵ်း	yut
ၵ်း	hu	ၵ်း	at	ၵ်း	yal	ၵ်း	yat	ၵ်း	wal	ၵ်း	hon
ၵ်း/ၵ်း	uu	ၵ်း	at	ၵ်း	yae	ၵ်း	yout	ၵ်း	wae	ၵ်း	honn
ၵ်း	ay	ၵ်း/ၵ်း	at	ၵ်း	yaw	ၵ်း	yite	ၵ်း	ha	ၵ်း	yun
ၵ်း	ae	ၵ်း	ate	ၵ်း	yut	ၵ်း	yin	ၵ်း	har	ၵ်း	hwa
ၵ်း	ay	ၵ်း	ote	ၵ်း	yaw	ၵ်း	hint	ၵ်း	ha	ၵ်း	lyar
ၵ်း	al	ၵ်း	hote	ၵ်း	yan	ၵ်း	yinn	ၵ်း	har	ၵ်း	haw
ၵ်း	ae	ၵ်း	hit	ၵ်း	yant	ၵ်း	yaung	ၵ်း	eet	ၵ်း	hawe
ၵ်း/ၵ်း	aww	ၵ်း	ot	ၵ်း	yo	ၵ်း	yout	ၵ်း	he	ၵ်း	haw
ၵ်း/ၵ်း	ot	ၵ်း	at	ၵ်း	yoe	ၵ်း	yaung	ၵ်း	hee	ၵ်း	hloe

ခေ/ခေ	aw	၁၁/၁၁	at	၁၁	yoe	၁၁	yai	၁၁	uu	၁၁	hlin
ခေ	ot	၁၁	kit	၁၁	hone	၁၁	yine	၁၁	hu	၁၁	hit
၁	an	၁၁	oat	၁၁	hote	၁၁	yit	၁၁	huu	၁၁	yinn
၁	ant	၁၁	ite	၁၁	yone	၁၁	yit	၁၁/၁၁	huu	၁၁	wa
၁/၁	o	၁၁	an	၁၁	yat	၁၁	yin	၁၁/၁၁	huu	၁၁	war
၁/၁	hoet	၁၁	ant	၁၁	yout	၁၁	yinn	၁၁	hay	၁၁	wyar
၁/၁	oe	၁၁	am	၁၁	yite	၁၁	yi	၁၁	hae	၁၁	way
၁/၁	one	၁၁	han	၁၁	at	၁၁	yae	၁၁	aye	၁၁	ywae
၁	ote	၁၁	ann	၁၁	yin	၁၁	yee	၁၁	hal	၁၁	way
၁	one	၁၁	ane	၁၁	yint	၁၁	yat	၁၁	hlae	၁၁	wyat
၁	hoe	၁၁	ate	၁၁	yin	၁၁	yat	၁၁	aww	၁၁	wyin
၁	at	၁၁	ane	၁၁	aung	၁၁	yate	၁၁	haw	၁၁	yoot
၁	ate	၁၁	one	၁၁	yaung	၁၁	yote	၁၁	han	၁၁	yoon
၁/၁	ote	၁၁	oat	၁၁	hai	၁၁	yan	၁၁	hant	၁၁	yut
၁/၁	ote	၁၁	oat	၁၁	hai	၁၁	yan	၁၁	hant	၁၁	yut
၁/၁	out	၁၁	one	၁၁	yite	၁၁	yant	၁၁	ho	၁၁	wyan
၁	ite	၁၁	at	၁၁	yine	၁၁	yan	၁၁	hoet	၁၁	ywal
၁	ite	၁၁	ate	၁၁	yit	၁၁	yai	၁၁/၁၁	hoe	၁၁	yu
၁	hote	၁၁	ote	၁၁	yin	၁၁	hyate	၁၁	hone	၁၁	yae
၁	at	၁၁	at	၁၁	yin	၁၁	yane	၁၁	hote	၁၁	yone
၁	ate	၁၁	at	၁၁	in	၁၁	yone	၁၁	hone	၁၁	yout
၁	oat	၁၁	ait	၁၁	yi	၁၁	yat	၁၁	hat	၁၁	yite
၁	in	၁၁	ote	၁၁	yeet	၁၁	yote	၁၁	hout	၁၁	yint
၁	int	၁၁	at	၁၁	yee	၁၁	yote	၁၁	hite	၁၁	yaung
၁	inn	၁၁	an	၁၁	yat	၁၁	yan	၁၁	hin	၁၁	yout
၁	ain	၁၁	amt	၁၁	yate	၁၁	yan	၁၁	hint	၁၁	yaung
၁/၁	aung	၁၁	eint	၁၁	yote	၁၁	yane	၁၁	hinn	၁၁	hote
၁/၁	out	၁၁	ann	၁၁	yan	၁၁	yate	၁၁	hane	၁၁	yote
၁/၁	aung	၁၁	ain	၁၁	ant	၁၁	yane	၁၁	haung	၁၁	yine
၁	ine	၁၁	aint	၁၁	yan	၁၁	yal	၁၁	hout	၁၁	lwa
၁	hite	၁၁	ane	၁၁	yane	၁၁	hlo	၁၁	haung	၁၁	hwar
၁	ine	၁၁	one	၁၁	yane	၁၁	at	၁၁	laing	၁၁	war
၁	aing	၁၁	hone	၁၁	yone	၁၁	hate	၁၁	laing	၁၁	wee
၁	it	၁၁	al	၁၁	yone	၁၁	at	၁၁	hit	၁၁	lwae
၁	it	၁၁	al	၁၁	yat	၁၁	at	၁၁	hin	၁၁	lwae
၁	hars	၁၁	ae	၁၁	yate	၁၁	ho	၁၁	hin	၁၁	wae
၁	ate	၁၁	hal	၁၁	yote	၁၁	wa	၁၁	lae	၁၁	wal
၁	hit	၁၁	o	၁၁	yote	၁၁	war	၁၁	hal	၁၁	won

ဝေ့	oss	ဝေ့	oe	ယံး	yan	ဝေ့	wa	ဟတ်	hat	ယံး	yon
ဝေ့	ost	ဝေ့	al	ယံး	aint	ဝေ့	warr	ဟတ်	hate	ယံး	hwin
ဝေ့	it	ဝေ့	an	ယံး	yal	ဝေ့	weet	ဟတ်	hote	ယံး	wint
ယံး	yiz	ဝေ့	hay	ယံး	ya	ဝေ့	weet	ဟတ်	han	ယံး	hwin
ဝေ့	ist	ဝေ့	ho	ယံး	yar	ဝေ့	wee	ဟတ်	lant	ယံး	oot
ယံး	yit	ဝေ့	Al	ယံး	yar	ဝေ့	hee	ဟတ်	han	ယံး	yon
ယံး	in	ဝေ့	ain	ယံး	yi	ဝေ့	way	ဟတ်	hane	ယံး	yon
ယံး	int	ဝေ့	ho	ယံး	ye	ဝေ့	wae	ဟတ်	hane	ယံး	wat
ယံး	inn	ဝေ့	all	ယံး	yee	ဝေ့	way	ဟတ်	hone	ယံး	wan
ယံး	im	ဝေ့	all	ယံး	yu	ဝေ့	wal	ဟတ်	hote	ယံး	yun
ယံး	yin	ဝေ့	ho	ယံး	yu	ဝေ့	wae	ဟတ်	hone		
ယံး	hone	ဝေ့	at	ယံး	yuu	ဝေ့	on	ဟတ်	hat		
ဝေ့	ay	ဝေ့	ate	ယံး	yay	ဝေ့	ont	ဟတ်	hate		
ယံး	i	ဝေ့	ote	ယံး	hae	ဝေ့	wat	ဟတ်	hote		
ယံး	eet	ဝေ့	aye	ယံး	way	ဝေ့	wite	ဟတ်	hann		
ယံး	ee	ဝေ့	ai	ယံး	yal	ဝေ့	win	ဟတ်	hane		
ဝေ့	oot	ယံး	ya	ယံး	hae	ဝေ့	wint	ဟတ်	ate		
ဝေ့	ate	ယံး	yar	ယံး	yaw	ဝေ့	win	ဟတ်	hal		

Warning: You may have noticed that “အ” is not romanised as “a”. This is because it functions as a silent, “placeholder” onset in spoken Burmese. Instead of typing, for example, “aa”, “aone”, “aote”, one would simply type “a”, “one”, or “ote”. This is consistent with how most Burmese type.

Bigger Warning: Please be careful not to read the rhymes with an English pronunciation. For example, when typing “one”, do *not* read it to sound like the English numeral 1. If you consider a consonant in front of the word (e.g., “kone”) then you won’t run into this problem.

4.1.2. Special Romanisations

Some words don’t fall into any simple pattern, and you’ll just have to memorize them. These include pat-sint and foreign words. Also, there are six words that directly contradict our onset + rhyme rule; these include “lhyo” (not “sho”) for example. The following table lists these six words, and the pat-sint words. For brevity, we do not include foreign words, which romanize to their foreign spelling.

Special Words							
Over-rides		မန္တ	manda	စဝ္ဂ	sitsa	စမ္မာ	sanpar
လျှို	lhyo	ဝန္တာ	wontar	အိတ္တ	itsa	သပ္ပိယ	thatnal
လျှိုး	lhyoe	ဂုတ္တ	goatta	ရိတ္တ	yitsa		
လျှက်	lhyat	သန္တာ	thandar	ကိတ္တ	kitsa	နန္ဒ	nanda
လျှင်	lhyin	သန္တာ	thanda	နိတ္တ	nitsa	စန္ဒ	sanda
လျှပ်	lhyat	အတ္တ	atta	မိတ္တ	myitzi	စန္ဒ	sanda
လျှမ်း	lhyan	လန္တတ်	lantate			အိန္ဒ	indi
စ		မုတ္တ	moteta	သိက္ခာ	theitkhao	ကိန္ဒ	kaida
ကမ္မာ	kaba	ရိတ္တ	yateta			အိဒြေ	eidray

ကမ္ဘာ	kamba	ဝ		ဒုက္ခ	dotka	အိန္ဒြာ	eidra
ကမ္ဘာ	kabar	ခန္ဓာ	khandar	ဒုက္ခိ	dotkhe	နန္ဒာ	nandar
ကမ္ဘာ	kambaw	ဝန္ဓု	bandu	ယက္ခ	yatkha	မုဒ္ဒ	moteda
မ		ဗုဒ္ဓ	buddha	ရက္ခ	ratka	ဣန္ဒြေ	eaidray
စက္ကန့်	setkant	ဝ		လက္ခ	latkha	ဝ	
စက္က	satku	ဓမ္မ	damma	သိက္ခာ	theitkhar	ဘဝ	batga
မက္က	matka	ဓမ္မာ	dammar	ရက္ခိ	yatkhe	ပုဂ္ဂိုလ်	poatgo
ကျိက္က	kyiteka	ကြမ္မာ	kyanmar	စက္က	satkhu	မဂ္ဂ	matga
ဥက္က	oatka	သမ္မာ	tanmar	အ		ပုဂ္ဂ	poatga
တက္က	tatka	သမ္မ	tanma	အင်္ဂ	inga	မ	
သက္က	thatka	ရမ္မတ်	yatmat	အင်္ဂါ	ingar	ပုဏ္ဏား	ponenar
မ		မ		ဒင်္ဂါး	dingar	မ	
မိလ္လာ	mailar	ဝတ္ထု	woothtu	နင်္ဂ	ninga	ဘဏ္ဍာ	bandar
မ		မိတ္တိ	matehti	မင်္ဂ	minga	သဏ္ဍာန်	thatan
သဏ္ဍာန်	thandan	မ		သင်္ကေ	tinkay	သဏ္ဍာန်	thandan
ပဏ္ဍာ	patnar	ပုဏ္ဏ	poatsar	ဒင်္ဂ	dinga	ဒဏ္ဍာ	dantar
ကဏ္ဍ	ganda	ဣတ္ထိ	eightsi	သင်္ကန်း	thingan	ကဏ္ဍ	ganda
မ		မိတ္တိ	myitsee	ဟင်္သာ	hintar	Misc	
စန္ဒ	santa	ရမ္မာန်	yutesan	နင်္ဂါ	ninlar	သုံ	the
အန္ဒ	anda	မ		ဘင်္ဂ	binga	လက်ျာ	latyar
နတ္ထိ	neththi	သင်္ချိုင်း	thingjine	သင်္ကြန်	thingyam	ဧည့်	ae
ကတ္တာ	kattar	မ		သင်္ဘော	thinbaw	ဧ	ae
ပြိတ္တာ	pyittar	အင်္ဂါ	inngi	သင်္ချာ	tinchar	ဤ	ei
မတ္တိ	matedhi	မ		စင်္ကာပူ	singapore	၏	eat
ဝတ္ထ	wootta	သဏ္ဍာ	tinzar	အင်္ဂလိပ်	english	၍	ywae
မေတ္တာ	myittar	ဝိဇ္ဇာ	waitzar	မ		၌	nhite
ခေတ္တ	khitta	မ		သိပ္ပံ	theitpan	ဦ	ou
ပုတ္တ	poatta	ပစ္စည်း	pyitsee	သဏ္ဍာ	thanpar	ဦး	oo
ပတ္တ	patta	သဏ္ဍာ	thitsar	သဏ္ဍာယ်	thetpal	၎င်း	lakaung
သတ္တ	tattar	ဥတ္တ	oatsar	ကပ္ပ	katpa	သာ	tha
သတ္တု	thattu	ဝစ္စ	wootsa	ကုမ္ပ	konepa	ယောက်ျား	youtkyar
သတ္တိ	thetthe	ပစ္စ	pyitsa	မိပ္ပိယ	datepal		
အန္တာ	antar	ကိစ္စ	katesa	စမ္မော	sanpaw		

Note: At the time of publication, the following three special last-minute exceptions were added:

Extra Exceptions	
အောင်	aww
အောင်	aww
အောင်	aww

4.1.3. Shortcuts

WaitZar uses two shortcuts to speed up typing.

- Any word containing “aung” can shorten it to simply “g”. For example, “kaung” can be typed as “kg”. “Aung” by itself can be shortened to “ag”.
- If only one set of words is possible at any point, the sequence to that set is parenthesised. For example, typing “singa” is sufficient for WaitZar to guess that you want “singapore”.

4.2. Dictionary Lookup

If, for any reason, you cannot find the word you want to type in WaitZar, you can search for it in the generated wordlist:

http://waitzar.googlecode.com/svn/trunk/eclipse_project/wordlist.generated.txt

In the future, WaitZar will allow you to search within the program itself; for now, you will have to manually look for your word. Please note that you will need Zawgyi-One installed to view the generated wordlist (most development is done in Zawgyi-One to avoid ambiguity errors).

5. Known Bugs

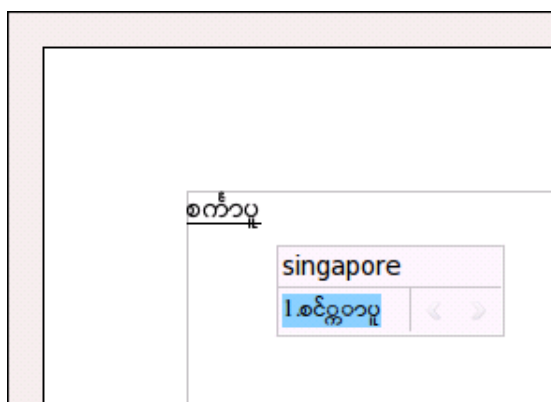
This section details a number of issues that are known to occur with WaitZar. If you find any new issues, please submit a bug report online:

<http://code.google.com/p/scim-waitzar/issues/entry/>

5.1. Pat-Sint Choices are Glitched

Even if you follow the advice in Section 2.4, you will find that some words are still incorrect when presented as choices. For example, below is a screenshot of “singapore” as it appears in the lookup table –note that the actual, typed version is still correct.

Figure 10. The Myanmar text is “decomposed” in SCIM, but renders correctly in Open Office.



The WaitZar team is uncertain as to what is causing this glitch. There are many interacting components (cairo, pango, GTK, or even scim itself) that might be slightly off. Perhaps the issue is actually a fontconfig issue on our end. We *could* solve this problem by defaulting SCIM to the Zawgyi-One font, but we feel that this solution would be a cheap hack that delays the adoption of Unicode 5.1 for Burmese in Linux. Instead, we will write some test cases to determine what package is causing the glitch, and then work with those developers to resolve this problem.

5.2. Alt+Tab Sometimes “Sticks”

This appears to be a bug in SCIM itself; under some configurations, switching to WaitZar will cause the “Alt+Tab” combination to “stick” –pressing this will outline the window one is hoping to switch to, but releasing it will see you at the original window; nothing happens. Clicking on any other window will cause this bug to go away until the next time you switch languages.

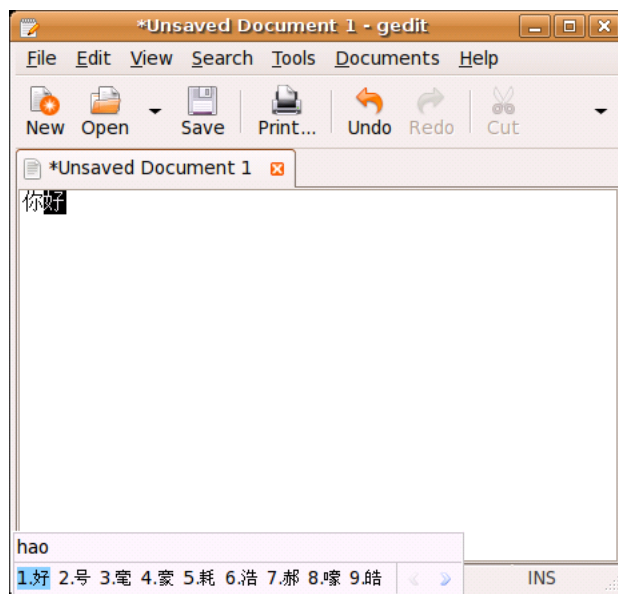
So far, we have only found one setup where this occurs, and this behavior is apparent in other IMEngines (Chinese, Thai, etc.), so it is probably a bizarre case of SCIM and Linux interacting poorly. Please report this bug if you encounter it, and we will try our best to narrow it down and fix it.

5.3. Choices Stuck in the Lower-Left

Under some configurations (usually, the same ones described in Section 5.2), the list of choices will appear “locked” to the lower-left of the current window. See the screenshot below for an example of this annoying bug, which makes

typing very difficult. Of course, you can click and drag the choices window to a more convenient location, but it will jump right back to the lower-left corner as soon as you switch the language again.

Figure 11. The choices window is “sticking” to the corner of gedit. This occurs with all language IMEs (Chinese is shown here).

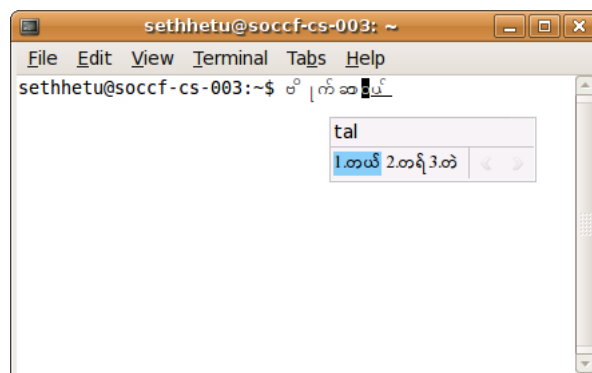


We might try re-positioning the choices window to somewhere nicer when WaitZar gets a “focus_in()” message, but seeing as how this is tied in to the previous bug, we decided not to hack around it until we were sure it couldn't be fixed normally.

5.4. The Console Looks Glitchy

Have a look at the following screenshot.

Figure 12. Notice how the text does not compose correctly in the console.



If this is what you see when you try to type WaitZar in the console, there's no need a file a bug report; we're aware. If you actually see proper Burmese, then *please* let us know how to configure the console to make this work. (**Note:** using Zawgyi-One or Win Innwa is *not* the solution we are looking for.)

A. License

scim-waitzar is licensed under the GNU General Public License v3, although some parts of it are licensed under the Apache License v2, and a small amount is licensed under the GNU General Public License v2.

This means that **scim-waitzar** is Open Source, of course, and free software as well. The original WaitZar on Windows is licensed under a BSD-style license; we chose to license **scim-waitzar** under the GPL to make it more compatible with the GNU philosophy, and chose version 3.0 to help strictly enforce that this software is not used in a non-open way.

Of course, our concerns apply only to Linux, really. The WaitZar library API, available from:

<http://code.google.com/p/waitzar/downloads/list>

...contains the majority of the encapsulated functionality; both the Windows *and* the Linux project build primarily off of this library, adding only implementation-specific code. Furthermore, the library is released under the Apache License, with the intention of not forcing any restrictions on developers.

Discussion about WaitZar is generally taken with both appreciation and seriousness, and occurs primarily on the mmgeeks forum, in English:

<http://mmgeeks.org/forum/>

B. How It Works

This section explains the use of trigram models in language modeling. It is currently under development.