Unicode Typography Primer

Roland Telfeyan
Charlotte, North Carolina
roland@telf.com

June 23, 2008

0 Introduction

If you would like to write in Greek, Coptic, Cyrillic, Armenian, Hebrew, Arabic, Chinese, and many other languages together in one document without fussing with fonts, if you want to type in any of these languages using your own personal keyboard layout, and if you want others to be able to read your text in absence of the font or keyboard layout or computer system you used, then you need Unicode.

Unicode is the international, de-facto standard which defines the characters and codes for all the world's type scripts. It is supported by all the major computer vendors.

1 Technical Background

1.1 Letters are Numbers

The text contained in a computer file is just an ordered list or array of numbers. If you located a plain text file on your computer's hard drive and looked at it with a microscope, you would see a sequence of numbers, each of which corresponds to a given letter. This article itself is nothing but a linear sequence of numbers. Depending upon the application program used, the numbers are interpreted and rendered as letters on your screen or your printer. The interpretation requires a definition of which letter is assigned to which number.

1.2 Unicode: A Union of Codes

The word "Unicode" is a contraction of two roots: "uni" and "code." First the "uni" part (from the Latin for "one, single"). Unicode is a *union* of all the world's alphabets into a single, giant world alphabet. Included are Latin, Greek, Coptic, Cyrillic, Armenian, Hebrew, Arabic, Syriac, Georgian, Ethiopic, and many others.

Now consider the "code" part. Unicode is also a *code* that establishes the fixed relation between letters and numbers that defines how letters are to be stored in computer files.[1] Unicode assigns a unique numerical code to each unique letter in the giant world alphabet. Latin 'n', Greek 'ν', Cyrillic 'н', Armenian 'ն', Hebrew 'נ', Arabic 'ن', Syriac 'ܢ', Georgian 'ნ', and Ethiopic 'ነ' all have the 'n' sound, but Unicode considers them to be different letters, each with its own unique numerical code. Across all the languages' alphabets, there are no two letters with the same code. At present there are more than 95,000 distinct alphabetic signs accounted for in the Unicode alphabet.

[1] Because all a computer's memory can hold is numbers, mappings and encodings and various other sorts of strucures are required to represent anything other than numbers.

1.3 Independence of Platform, Application, and Font

By mapping a unique number for each world alphabet letter, it is possible to have a plain text file which contains many language alphabets together without using any particular font or style or application or computer platform or any other feature to differentiate the alphabets. Unicode therefore makes it possible for people using different applications on different computer systems to send and receive text in any alphabet without having to specify fonts. The Unicode text definition is presently well-supported by Macintosh, Windows, and Linux operating systems.

2 Historical Background

A little bit of historical background should illustrate some of the challenges associated with multi-lingual typography that Unicode addresses and answers.

2.1 Early 1960's: ASCII

In 1963, when the ASCII (American Standard Code for Information Interchange) standard was invented, no more than 128 different codes and alphabetic signs were planned for. ASCII was designed to specify the encoding for teletype messages.

```
  0 nul    1 soh    2 stx    3 etx    4 eot    5 enq    6 ack    7 bel
  8 bs     9 ht    10 nl    11 vt    12 np    13 cr    14 so    15 si
 16 dle   17 dc1   18 dc2   19 dc3   20 dc4   21 nak   22 syn   23 etb
 24 can   25 em    26 sub   27 esc   28 fs    29 gs    30 rs    31 us
 32 sp    33  !    34  "    35  #    36  $    37  %    38  &    39  '
 40  (    41  )    42  *    43  +    44  ,    45  -    46  .    47  /
 48  0    49  1    50  2    51  3    52  4    53  5    54  6    55  7
 56  8    57  9    58  :    59  ;    60  <    61  =    62  >    63  ?
 64  @    65  A    66  B    67  C    68  D    69  E    70  F    71  G
 72  H    73  I    74  J    75  K    76  L    77  M    78  N    79  O
 80  P    81  Q    82  R    83  S    84  T    85  U    86  V    87  W
 88  X    89  Y    90  Z    91  [    92  \    93  ]    94  ^    95  _
 96  `    97  a    98  b    99  c   100  d   101  e   102  f   103  g
104  h   105  i   106  j   107  k   108  l   109  m   110  n   111  o
112  p   113  q   114  r   115  s   116  t   117  u   118  v   119  w
120  x   121  y   122  z   123  {   124  |   125  }   126  ~   127 del
```

Figure 1. The ASCII Character Set

ASCII represented printed characters like 'A', 'B', '+', '=', etc., and it included commands to control the print head of the teletype, like "carriage return", "line feed", "tab", "back space", etc. Like Unicode, it gave no indication of typeface appearance, just numbers representing letters. Unlike Unicode, it defined only 128 characters (Unicode started by defining tens of thousands). The limited ASCII standard became embedded at the core of every computer system and was used for a long time. Its entrenchment eventually stunted the growth of multi-lingual computing.

2.2 Mid-1970's: Computer Fonts

Early computers used the ASCII encoding to store text in memory without paying any attention to graphics or type styles. Keyboards were designed physically to save characters directly into files. If you pressed an uppercase 'C' on the keyboard, the number 67 was stored in the file (Figure 2).

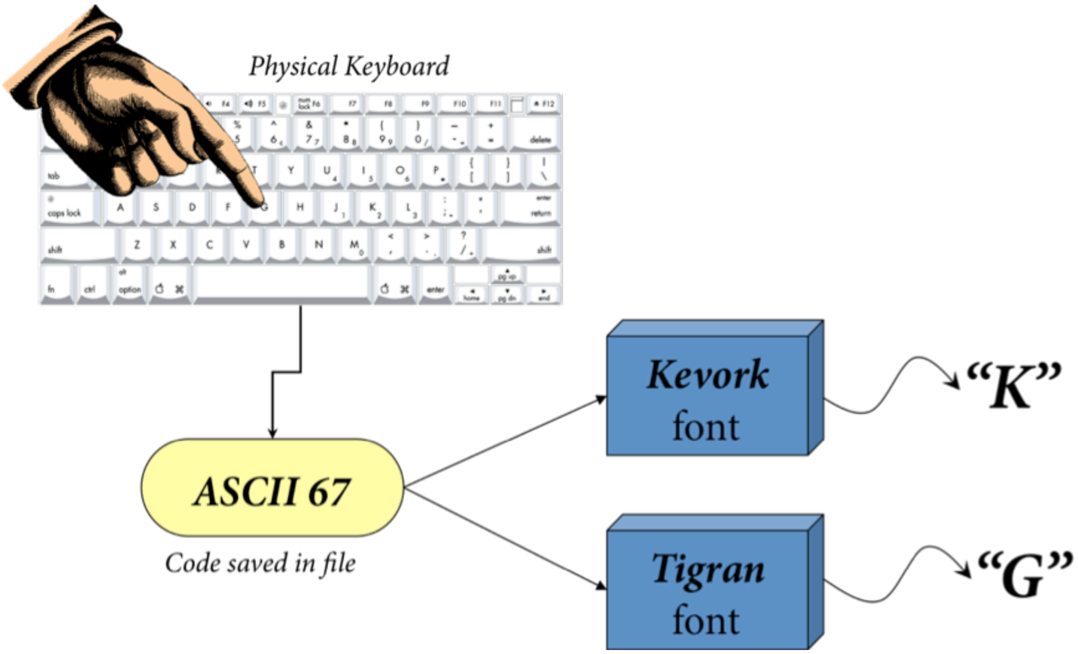


Figure 2. In 1985, fonts allowed multiple graphic representations of the same character key code, but there was no definitive character encoding standard, so the same character key code could look like anything from one font to another.

People desired different graphical representations for letters. Fonts fulfilled that desire by creating an association between ASCII codes and graphical representations, adding a typographical dimension to computers. The computer evolved from typewriter to typesetter. [2]

With fonts, the keyboard would still emit an ASCII character code into the text file, but now the computer would also store with that ASCII character a reference to a particular font or type style, such as Times or Helvetica, chosen by the user. If the user changed the font, the same text codes would take on a different appearance. The character code 97, say, could be rendered variously as ‘a’, ‘*a*’, ‘***a***’, or for that matter ‘**α**’, ‘**ա**’, ‘ა’, etc.

Multi-lingual computing was thus born. Fonts were the first flexible mapping interposed between the hardware keyboard and the printed signs.

2.3 Font Frustrations

Keyboards were physically designed to emit ASCII characters. Font designers suited personal preferences for keyboard layouts when assigning alphabetic signs to character codes. This meant that each font had its own ordering of letters that did not correspond to the consecutive lexical order of the alphabet. As a result, Armenian text could neither be sorted in a database or spreadsheet and even plain text could not be shared readily or viewed reliably in the absence of the very font used to create it.[3]

[2] Font technology pioneered at Xerox PARC was made popular by the Apple Macintosh computer system.

[3] To illustrate, suppose you were going to create an Armenian font in the 1980’s. Most font designers did not do the simple thing and assign ‘**ա**’ to 97, ‘**բ**’ to 98, ‘**գ**’ to 99, etc. because this type of arrangement would produce an

Another frustration was that the earliest fonts worked only on Macintosh systems, and Windows fonts, when they appeared, did not work on Macintosh systems. Therefore, sharing of text files could only occur between two people with the same operating system, same fonts, and same application programs.

2.4 1990’s: Keyboard Layouts

A company called NeXT Computer[4] fought font frustrations by implementing user-definable keyboard layouts as early as 1990. This technology started appearing on Macintosh systems in 2000. The placement of letters on a keyboard became a user preference, like the location of windows on a screen.

Keyboard layouts simplified the design requirements of fonts—eliminating the question of where the alphabetic signs would end up on the keyboard. Keyboard layouts were a second layer of mapping interposed between the hardware keyboard and the printed glyphs. They made the way clear for fonts to employ lexically-ordered encodings.

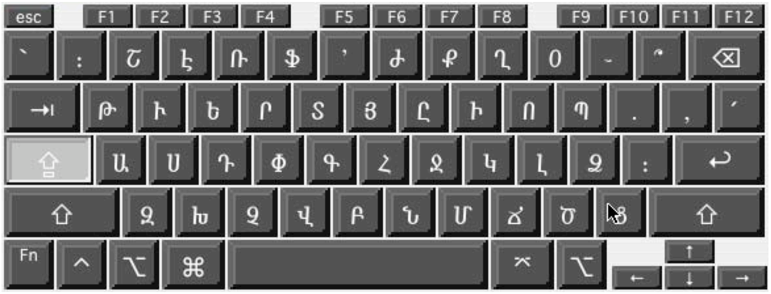


Figure 3. Keyboard layouts become separate from fonts, paving the way for character encoding standards.

2.5 Armenian Encodings

Armenians and Armenologists created dozens of different font encodings. Some were lexically ordered and used keyboard mappings, others were very plain and were not lexically ordered. In any case, no mainstream computer supplier ever supported any of them. Different Armenian fonts and encodings were used all around the world, with very little compatibility between them.

unintelligible keyboard layout. People sought phonetic layouts that mirrored the Latin typewriter layout. Start with the letter ‘ա’. If you assign it ASCII code 97, then the user will have to press ‘a’ on the keyboard to get ‘ա’. (However, if you want to implement a Royal typewriter layout, you would assign ‘ա’ to ASCII code 103, so that the user gets it by pressing ‘g’.) Now take ‘բ’. Does it go on 98 (‘b’) or 112 (‘p’) (or 101 ‘e’ for the Royal layout). You quickly see the quagmire produced here. No two fonts are compatible, and encoding Armenian letters based on keyboard layout meant that if you try to type some names and sort them in a spreadsheet, they will certainly be all out of order.

[4] When Steve Jobs was ousted from Apple in 1985, he pulled people out of Apple and started a new computer company called NeXT. Jobs (whose mother is a Hagopian) recruited Avadis Tevanian out of Carnegie Mellon University to create the NeXT computer operating system, called Mach. By 2000, Apple had bought NeXT for $400 million, and Jobs and Tevanian had taken command of Apple. Today everything going on at Apple with regard to hardware and software is 90% technology invented decades before, either at NeXT or at XEROX Palo Alto Research Center (PARC) which over thirty years ago foresaw so many technologies now viewed as standard, mouse, graphics, fonts, copy/paste, networked files, Ethernet, and many others.

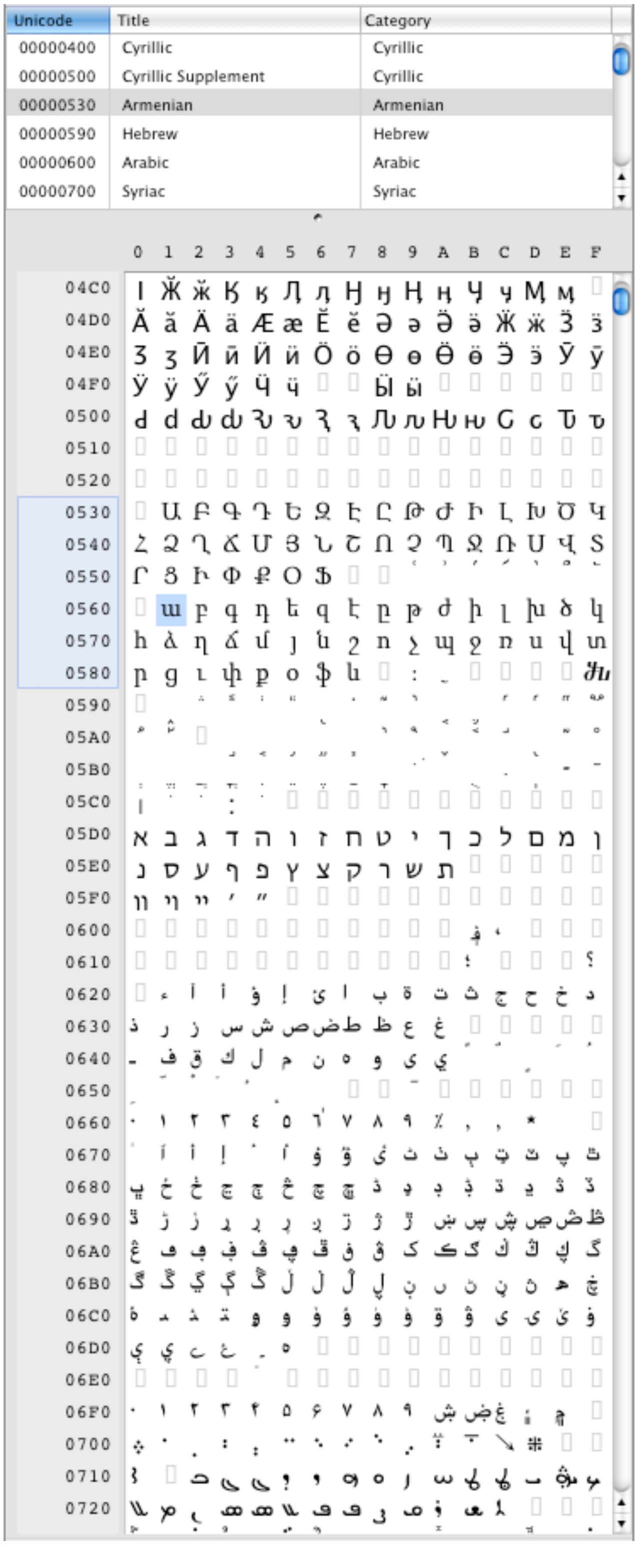
Because of the plethora of fonts and encoding "standards," digital texts were likely to become gibberish when transported. A scholar at St. Nersess Seminary would use custom-made Armenian PostScript fonts[5] for the Macintosh platform. A researcher at the Matenadaran would use ARMSCII-8[6] Windows TrueType fonts. The two were completely incompatible, and documents sent from one to the other would be totally illegible.

2.6 Font Compatibility

In the 1990s, Adobe, Apple, and Microsoft agreed and cooperated on new font specifications called TrueType and OpenType. By the year 2000, nobody was talking about Windows or Macintosh fonts anymore. In addition, the new OpenType specification allowed for fonts that could hold tens of thousands of characters, paving the way for the Unicode specification which had already been forming.

2.7 2000: Enter Unicode

The Unicode Character Set is a standard definition of character codes for the glyphs of most known languages. It is a product of the Unicode Consortium of computer vendors and language scholars. Whereas ASCII defined 128 characters, Unicode allows for more than 95,000 characters. Armenian codes range from 1328 to 1423 (95 codes). The goal of Unicode is to define a set of standard character codes for every known language.



Unicode	Title	Category
00000400	Cyrillic	Cyrillic
00000500	Cyrillic Supplement	Cyrillic
00000530	Armenian	Armenian
00000590	Hebrew	Hebrew
00000600	Arabic	Arabic
00000700	Syriac	Syriac

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
04C0	Ӏ	Ӂ	ӂ	Ӄ	ӄ	Ӆ	ӆ	Ӈ	ӈ	Ӊ	ӊ	Ӌ	ӌ	Ӎ	ӎ	□
04D0	Ӑ	ӑ	Ӓ	ӓ	Ӕ	ӕ	Ӗ	ӗ	Ә	ә	Ӛ	ӛ	Ӝ	ӝ	Ӟ	ӟ
04E0	Ӡ	ӡ	Ӣ	ӣ	Ӥ	ӥ	Ӧ	ӧ	Ө	ө	Ӫ	ӫ	Ӭ	ӭ	Ӯ	ӯ
04F0	Ӱ	ӱ	Ӳ	ӳ	Ӵ	ӵ	□	□	Ӹ	ӹ	□	□	□	□	□	□
0500	Ԁ	ԁ	Ԃ	ԃ	Ԅ	ԅ	Ԇ	ԇ	Ԉ	ԉ	Ԋ	ԋ	Ԍ	ԍ	Ԏ	ԏ
0510	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0520	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0530	□	Ա	Բ	Գ	Դ	Ե	Զ	Է	Ը	Թ	Ժ	Ի	Լ	Խ	Ծ	Կ
0540	Հ	Ձ	Ղ	Ճ	Մ	Յ	Ն	Շ	Ո	Չ	Պ	Ջ	Ռ	Ս	Վ	Տ
0550	Ր	Ց	Ւ	Փ	Ք	Օ	Ֆ	□	□	ՙ	՚	՛	՜	՝	՞	՟
0560	□	ա	բ	գ	դ	ե	զ	է	ը	թ	ժ	ի	լ	խ	ծ	կ
0570	հ	ձ	ղ	ճ	մ	յ	ն	շ	ո	չ	պ	ջ	ռ	ս	վ	տ
0580	ր	ց	ւ	փ	ք	օ	ֆ	և	□	։	֊	□	□	□	□	□
05D0	א	ב	ג	ד	ה	ו	ז	ח	ט	י	ך	כ	ל	ם	מ	ן
05E0	נ	ס	ע	ף	פ	ץ	צ	ק	ר	ש	ת	□	□	□	□	□
05F0	װ	ױ	ײ	׳	״	□	□	□	□	□	□	□	□	□	□	□

Figure 1. A very small segment of the Unicode character set showing Cyrillic, Armenian, Hebrew, Arabic, and Syriac.

[5] Lines, Fonts & Circles Armenian PostScript fonts were developed by Vazgen Aghajani.

[6] ARMSCII-8 was developed in Armenia. See **http://en.wikipedia.org/wiki/ARMSCII** for further information.

For the first time, Armenian (and many other) alphabets had their own individual codes defined by an international consortium of computer vendors and language scholars. Today all modern Mac and PC computers include Unicode of which Armenian is a part. Unicode has promoted and facilitated the compatibility between Apple and PC computers. For example, this article is a Word document that looks identical on Mac and Windows, without paying any attention to fonts.

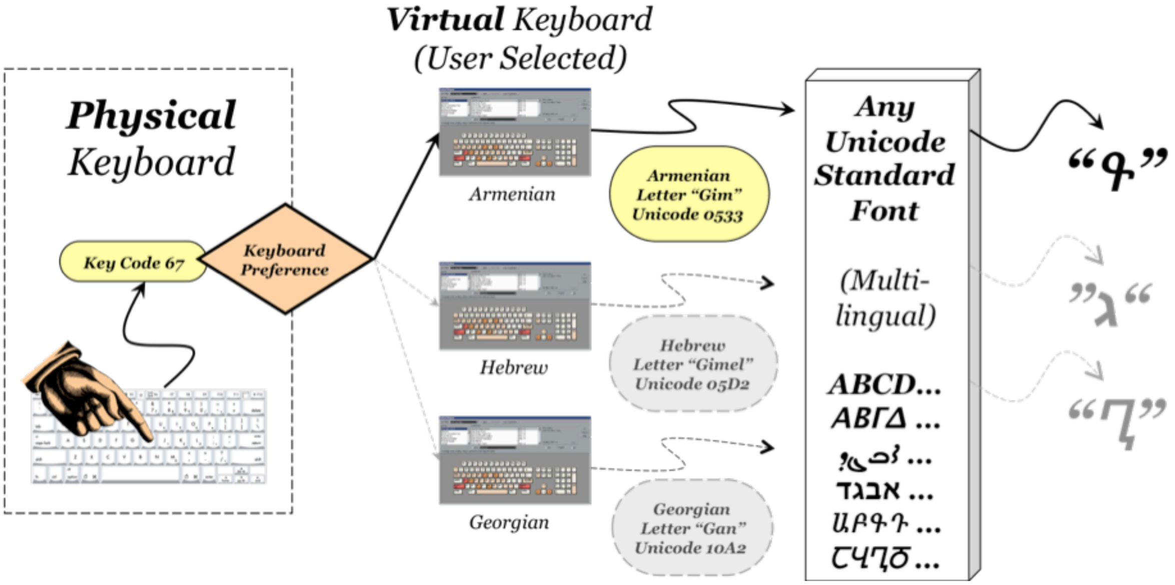


Figure 4. How Unicode works. The user presses the 'G' key. The computer sends the 'G' key to the currently selected keyboard layout which at the moment happens to be Armenian. The keyboard layout determines that the 'G' key corresponds to Armenian letter գ which has a character code of 0533. Any standard Unicode font that implements the Armenian letter գ glyph must implement that glyph at location 0533.

Unicode had been a concept in 1987 at Xerox, became a consortium of computer companies in 1991, and came into its full usefulness and practicability in the Windows 2000 and Macintosh 9 operating systems. Those systems presented some difficulties, but with the further developments of Mac OS X Tiger and Leopard and Windows XP and Vista the use of Unicode has become simple.

3 Using Unicode on Macintosh

Unless you are using an antique, Unicode is on your computer, from Windows XP and Mac OS X on. Here is a practical guide to the use of Unicode. This guide is oriented to the Macintosh, but the steps are analogous for Windows.

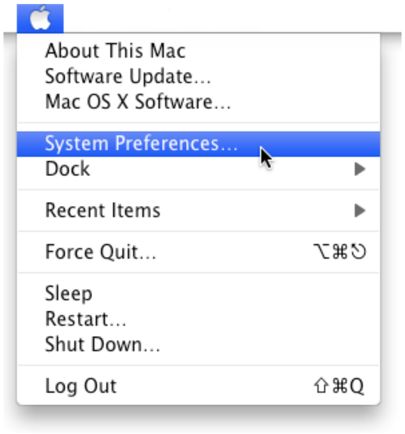
3.1 Choice of Input Methods

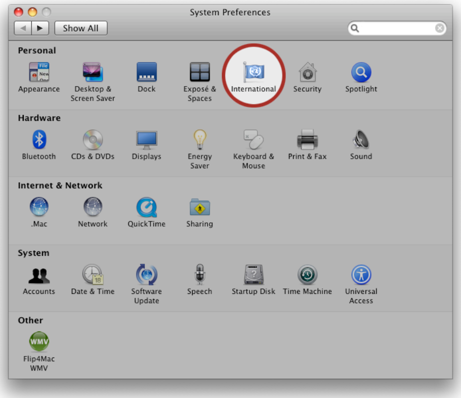
There are three types of input methods:

- A palette of characters that you click on with the mouse
- A mini on-screen software keyboard viewer that you click on with the mouse
- An alphabet and keyboard layout that defines what happens when you type

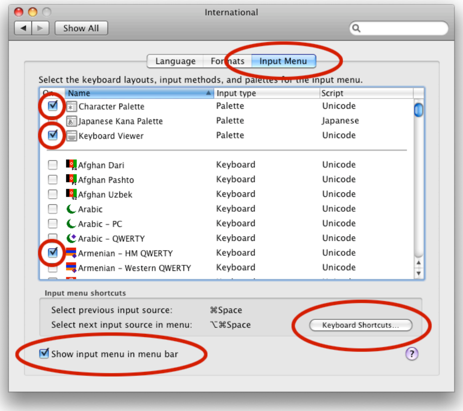
If you have not already done so, you should enable the **Character Palette**, the **Keyboard Viewer**, and any number of alphabet-keyboard-layout **Input Methods**. Choose ones that you will work with most often. In this example, we show how to choose Armenian HM QWERTY, but at any time you can come back and add more. Here are the steps:

Click on the Apple menu, choose **System Preferences**, and then click on **International**.





Select the **Input Menu** tab. Check **Character Palette**, **Keyboard Viewer**, and **Armenian – HM QWERTY**. Here you are selecting an alphabet and a keyboard for that alphabet. Also be sure that **Show input menu in menu bar** is checked.



Select the keyboard layouts, input methods, and palettes for the input menu.

On	Name	Input type	Script
☑	Character Palette	Palette	Unicode
□	Japanese Kana Palette	Palette	Japanese
☑	Keyboard Viewer	Palette	Unicode
□	Afghan Dari	Keyboard	Unicode
□	Afghan Pashto	Keyboard	Unicode
□	Afghan Uzbek	Keyboard	Unicode
□	Arabic	Keyboard	Unicode
□	Arabic – PC	Keyboard	Unicode
□	Arabic – QWERTY	Keyboard	Unicode
☑	Armenian – HM QWERTY	Keyboard	Unicode
□	Armenian – Western QWERTY	Keyboard	Unicode

Input menu shortcuts

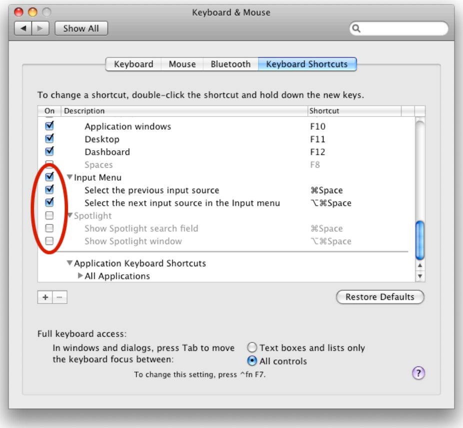
Select previous input source: ⌘Space

Select next input source in menu: ⌥⌘Space

☑ Show input menu in menu bar

Note that the selection of an alphabet is tightly connected to the selection of the keyboard layout for that alphabet. It is possible to have multiple keyboard layouts for the same alphabet; for example, **Armenian – Western QWERTY** is another keyboard layout for the Armenian alphabet.

Click the **Keyboard Mapping** button, and uncheck the **Spotlight** check boxes, and check the **Input Menu** check boxes. This setting enables you to switch your keyboard from one language to another by simply pressing **command-spacebar**.



To change a shortcut, double-click the shortcut and hold down the new keys.

On	Description	Shortcut
☑	Application windows	F10
☑	Desktop	F11
☑	Dashboard	F12
□	Spaces	F8
☑	▼Input Menu	
☑	Select the previous input source	⌘Space
☑	Select the next input source in the Input menu	⌥⌘Space
□	▼Spotlight	
□	Show Spotlight search field	⌘Space
□	Show Spotlight window	⌥⌘Space
	▼Application Keyboard Shortcuts	
	▶All Applications	

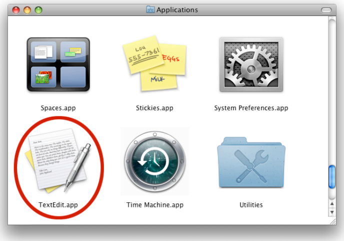
Full keyboard access:

In windows and dialogs, press Tab to move the keyboard focus between: ○ Text boxes and lists only ● All controls

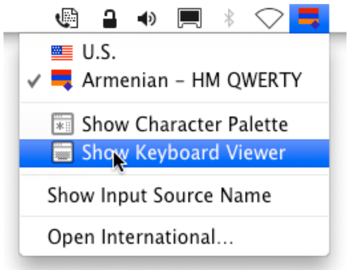
To change this setting, press ^fn F7.

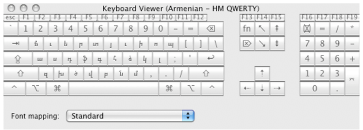
3.2 Input by Typing

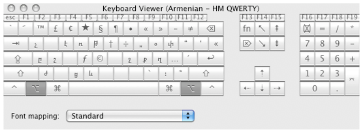
Open TextEdit from the Applications folder. TextEdit will open an empty window.



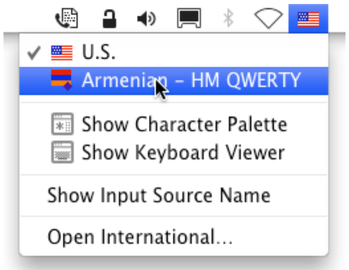
Open the **Keyboard Viewer** to learn where the letters are.

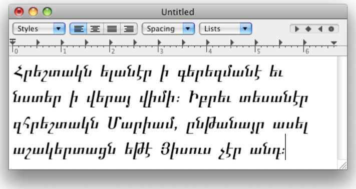






Press **command-spacebar** or, alternatively, choose the **Armenian** Language from the International Input menu in the menu bar, and then start typing.



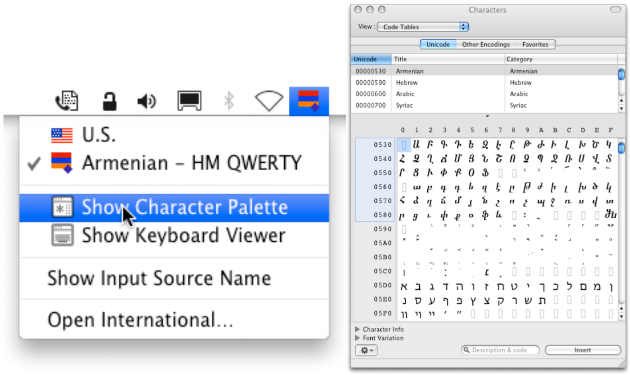


Հրեշտակն ելանէր ի գերեզմանէ եւ նստեր ի վերայ վիմի։ Իբրեւ տեսանէր զհրեշտակն Մարիամ, ընթանայր ասել աշակերտացն եթէ Յիսուս չէր անդ։

3.3 Input from the Character Palette

The Keyboard Viewer may not satisfy every need to locate a particular character. In this case, open the Character Palette to see all the letters of the entire Unicode character set in alphabetical order. Letters are grouped by "code pages." Each code page is a single writing script belonging to one language family. There are some code pages that contain nothing but mathematical or other symbols.

The Character Palette is useful for selecting a limited number of characters that may not be easily found on any keyboard. When you select a certain character, it shows you what that character looks like in all the different fonts installed on your system. You can place the desired character from the desired font into your document.

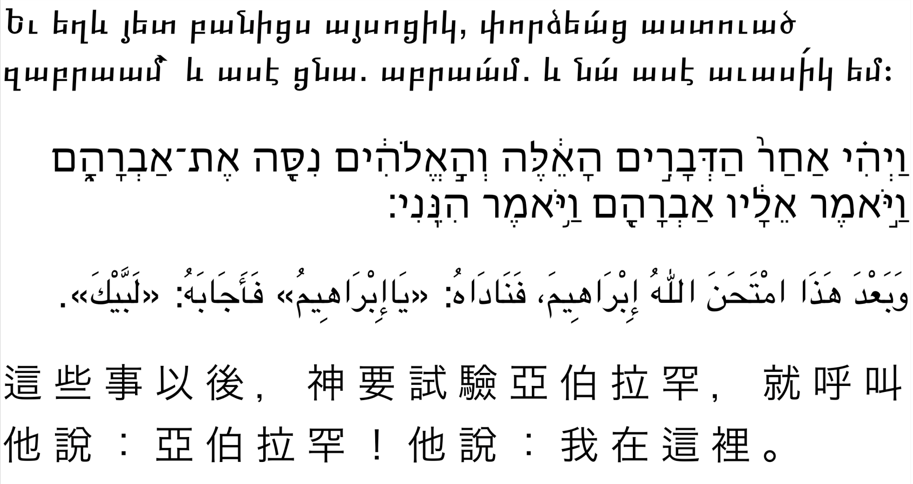


Unicode	Title	Category
00000530	Armenian	Armenian
00000590	Hebrew	Hebrew
00000600	Arabic	Arabic
00000700	Syriac	Syriac

Figure 5. The Character Palette helps you find any arbitrary character in the Unicode character set, regardless of its potential placement on any keyboard layout.

3.4 Mixing Multiple Languages

In order to enter multiple language alphabets in the same document, all you have to do is select the language from the International Input menu and start typing. To rapidly switch between languages, press **command-spacebar**. You do not have to set any font unless you require a certain typeface style.



Եւ եղև յետ բանիցս այսոցիկ, փորձեա՛ց աստուած
զաբրահա՞մ և ասէ ցնա. աբրահա՛մ. և նա՛ ասէ աւասի՛կ եմ:

וַיְהִ֗י אַחַר֙ הַדְּבָרִ֣ים הָאֵ֔לֶּה וְהָ֣אֱלֹהִ֔ים נִסָּ֖ה אֶת־אַבְרָהָ֑ם
וַיֹּ֣אמֶר אֵלָ֔יו אַבְרָהָ֖ם וַיֹּ֥אמֶר הִנֵּֽנִי׃

وَبَعْدَ هَذَا امْتَحَنَ اللّٰهُ إِبْرَاهِيمَ، فَنَادَاهُ: «يَاإِبْرَاهِيمُ» فَأَجَابَهُ: «لَبَّيْكَ».

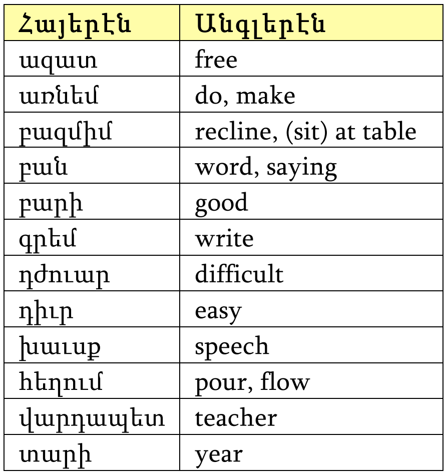
這些事以後，神要試驗亞伯拉罕，就呼叫
他說：亞伯拉罕！他說：我在這裡。

Figure 6. Apple’s basic TextEdit application (roughly the equivalent of WordPad) automatically writes from right to left when you choose Hebrew or Arabic.

3.5 Lexical Ordering

In the past, fonts that were arranged with keyboard layout in mind were unusable for the purpose of sorting lists of information. In Unicode, characters are all lexically ordered because keyboard layouts are a separate user preference.

Dealing with ordered lists of multi-lingual data in Excel or in a database is extremely straightforward when using Unicode. Here is a table of words sorted using Excel.



Հայերէն	Անգլերէն
ազատ	free
առնեմ	do, make
բազմիմ	recline, (sit) at table
բան	word, saying
բարի	good
գրեմ	write
դժուար	difficult
դիւր	easy
խաւսք	speech
հեղում	pour, flow
վարդապետ	teacher
տարի	year

Figure 7. A list of words is typed in Excel and then sorted by pressing the Sort A➔Z button.

3.6 Creating Keyboard Layouts

A keyboard layout can contain characters from anywhere in the Unicode character set. Generally keyboard layouts are confined to one particular alphabet, but it is possible to create one which incorporates characters from any number of alphabets as long as there are enough keys to go around.

For information on creating your own custom keyboard layout on either Macintosh or Windows, see **http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&cat_id=InputResources**.

3.7 Better Armenian Keyboard Layouts

The commercially-available Armenian keyboard layouts built into Mac OS X and Windows XP require the user to press the alt or option key in order to access the roughly ten Armenian characters that do not fit on top of the Latin alphabetic keys.

Noting that numbers are always available, either on the calculator keypad or by switching keyboards layouts, I decided to create an Armenian keyboard layout that exposes all characters without the need for the alt or option key. The layout **Armenian-RT** is freely available from **telf.com** (then click on **Keyboards**). Also available is **Armenian-Olympia** created by Hovhannes Kizoghian of Gandzasar.

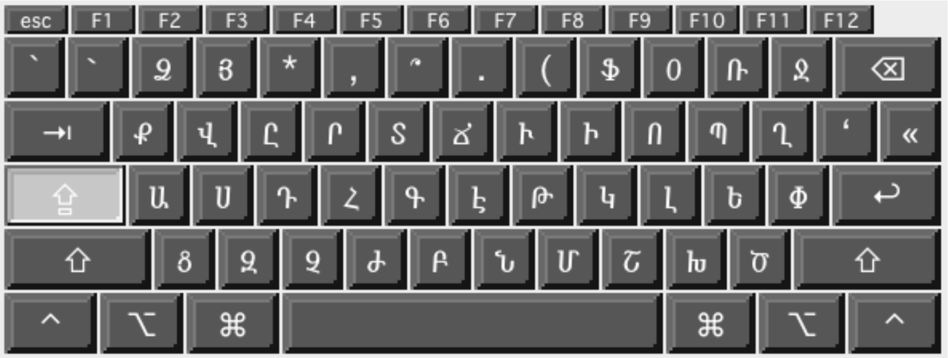

esc F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12
` ` Ձ Յ * , ՞ . (Ֆ Օ Ռ Ջ ⌫
→| Ք Վ Ը Ր Տ Ճ Ւ Ի Ո Պ Ղ ʻ «
⇪ Ա Ս Դ Հ Գ Է Թ Կ Լ Ե Փ ↩
⇧ Ց Ջ Ձ Ժ Բ Ն Մ Շ Խ Ծ ⇧
^ ⌥ ⌘ ⌘ ⌥ ^

Figure 8. Armenian-Olympia keyboard layout.

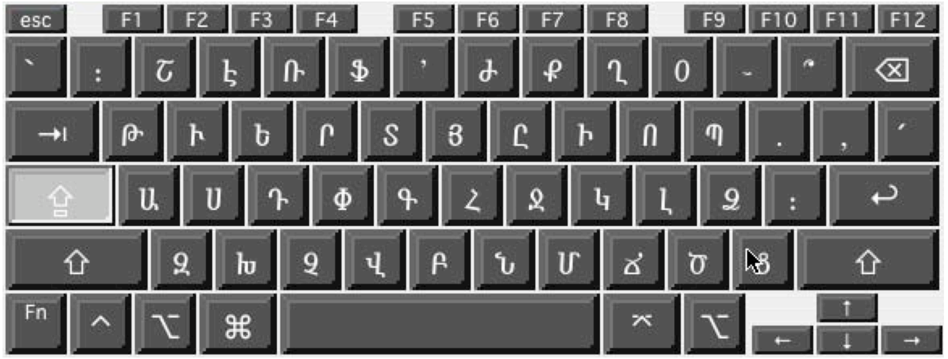

esc F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12
` : Շ Է Ռ Ֆ ՚ Ժ Ք Ղ Օ - ՞ ⌫
→| Թ Ւ Ե Ր Տ Յ Ը Ի Ո Պ . , ՛
⇪ Ա Ս Դ Փ Գ Հ Ջ Կ Լ Ձ ։ ↩
⇧ Զ Խ Ձ Վ Բ Ն Մ Ճ Ծ Ց ⇧
Fn ^ ⌥ ⌘ ⌅ ⌥ ↑ ← ↓ →

Figure 9. Armenian-RT keyboard layout. Both these keyboard layouts expose all Armenian characters without having to press the alt or option key. Both are free at telf.com.

3.8 TR: Converting Old Texts to Unicode

All modern computers today support Unicode well. However there is a profusion of ASCII and ARMSCII text data that scholars have accumulated over decades, and this data is illegible via Unicode.

I created a program called **TR** that makes extremely short work of converting any body of text to Unicode. It can convert text created with a dozen different input systems to Unicode text. Creating translators for **TR** takes a few seconds and anyone can do it without having to be a programmer. For a free download and instructions on how to use **TR**, see the **telf.com** website.

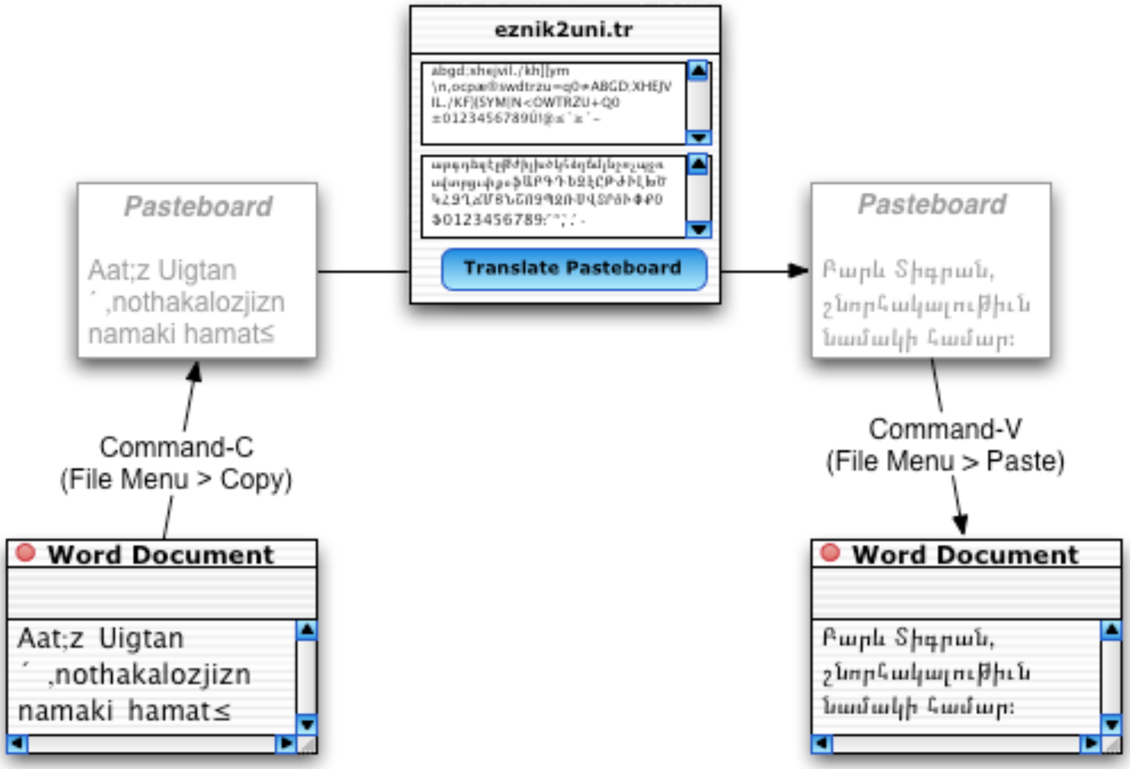


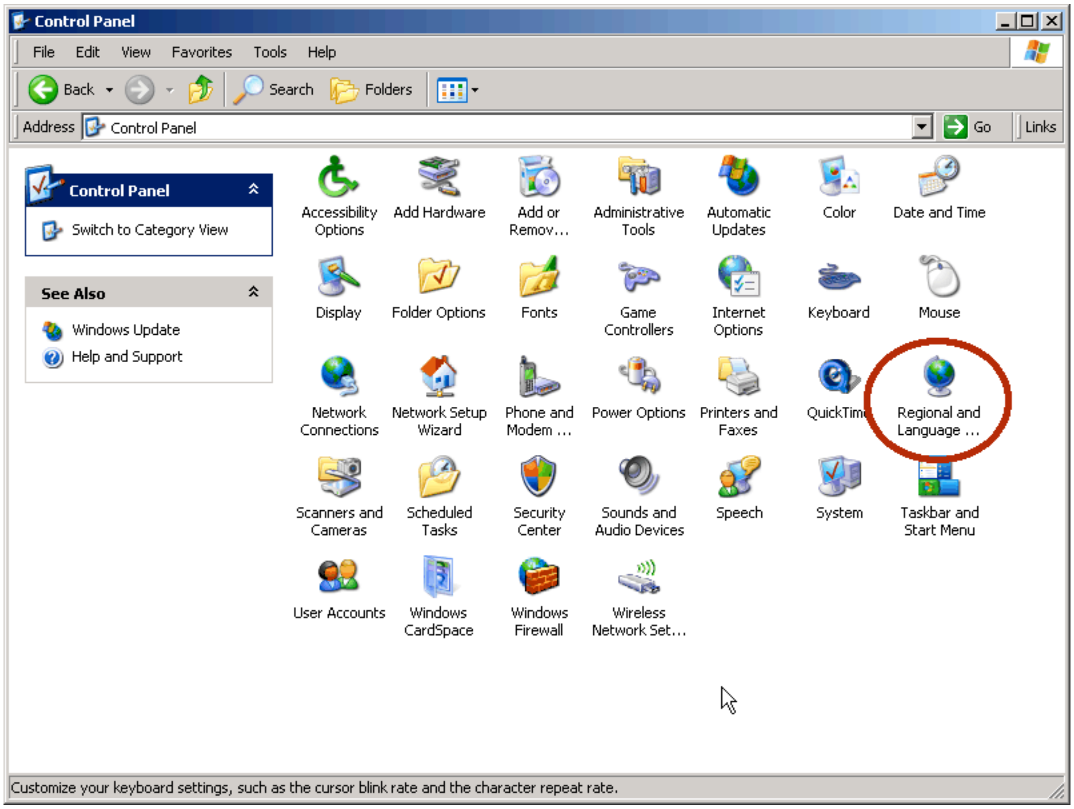
Figure 10. TR is a character encoding translation program available from telf.com. It understands a dozen different character encodings and converts text from one encoding to another in seconds. A customized encoding translator can be created in minutes simply by entering two alphabets.

4 Using Unicode on Windows

Setting up Unicode on Windows has the same basic steps as on the Macintosh: enabling the languages and keyboard layouts and switching between keyboard layouts while typing.

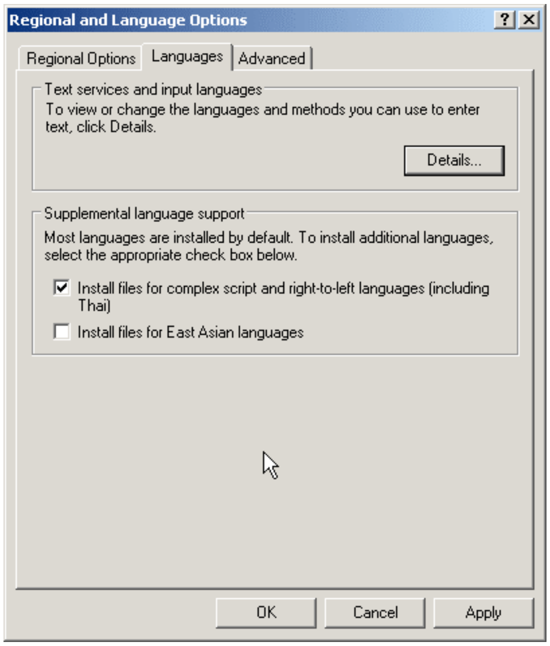
4.1 Input Methods

On Windows, in order to enable Armenian, you have to enable Asian languages. Go to the control panel and open the **Regional and Language** preferences.



Customize your keyboard settings, such as the cursor blink rate and the character repeat rate.

Check the box **Install files for complex script and right-to-left languages (including Thai)** and press the **OK** button.



Regional and Language Options

Text services and input languages

To view or change the languages and methods you can use to enter text, click Details.

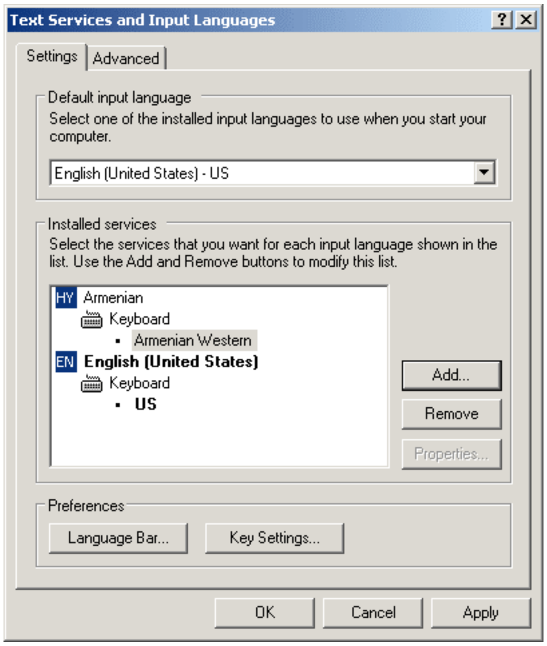
Supplemental language support

Most languages are installed by default. To install additional languages, select the appropriate check box below.

☑ Install files for complex script and right-to-left languages (including Thai)

□ Install files for East Asian languages

Now press the **Details…** button.



Text Services and Input Languages

Default input language

Select one of the installed input languages to use when you start your computer.

English (United States) - US

Installed services

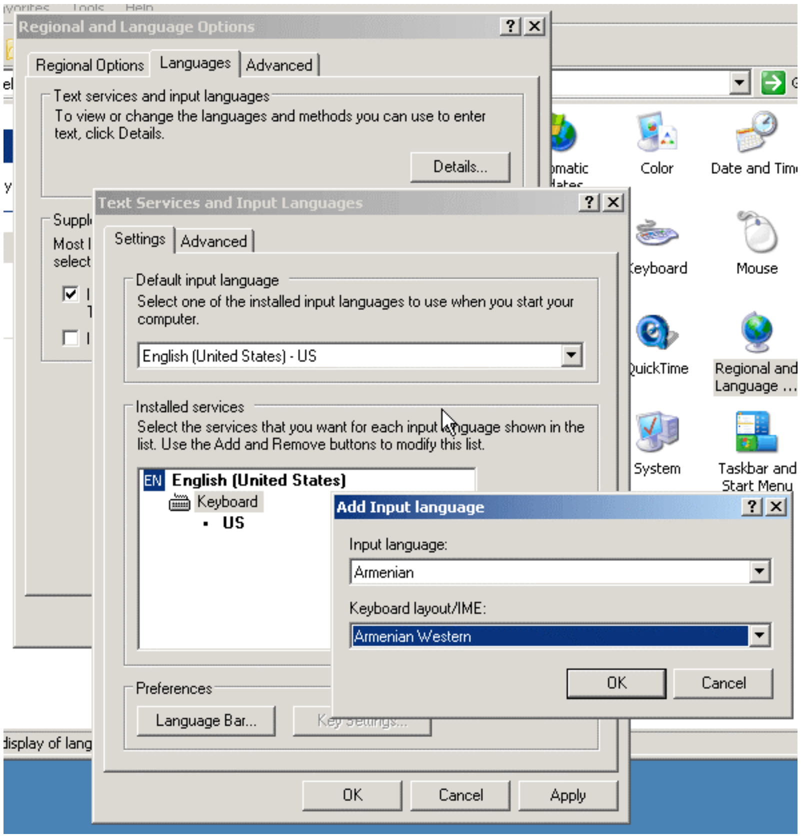
Select the services that you want for each input language shown in the list. Use the Add and Remove buttons to modify this list.

HY Armenian
- Keyboard
  - Armenian Western

EN **English (United States)**
- Keyboard
  - **US**

Preferences

Press the **Add…** button, and add Armenian as an input language and select an Armenian keyboard layout.



Text services and input languages

To view or change the languages and methods you can use to enter text, click Details.

Default input language

Select one of the installed input languages to use when you start your computer.

English (United States) - US

Installed services

Select the services that you want for each input language shown in the list. Use the Add and Remove buttons to modify this list.

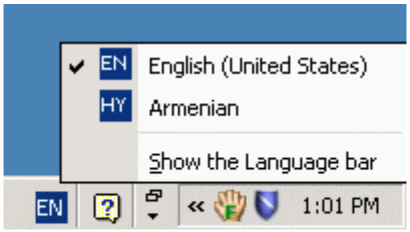
EN English (United States)
- Keyboard
  - US

Input language: Armenian

Keyboard layout/IME: Armenian Western

4.2 Typing

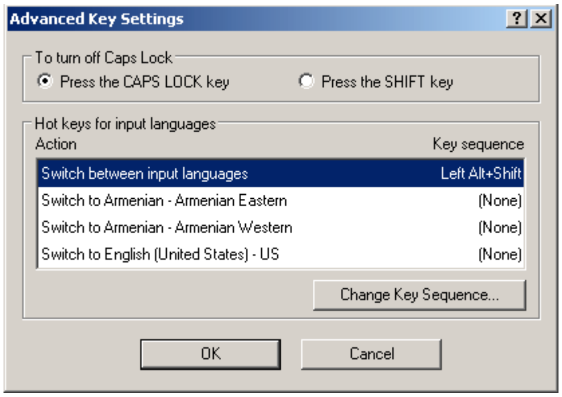
In the Windows Task Bar, use the mouse to select the different keyboards that are now enabled.



Open any text window (Word, Wordpad, Notepad) and start typing.

4.3 Custom Hot Keys for Keyboard Switching

While viewing the **Text Services and Input Languages** window, press the **Key Settings…** button. You may use this interface to create your own “hot keys” for switching between keyboards.



Advanced Key Settings

To turn off Caps Lock

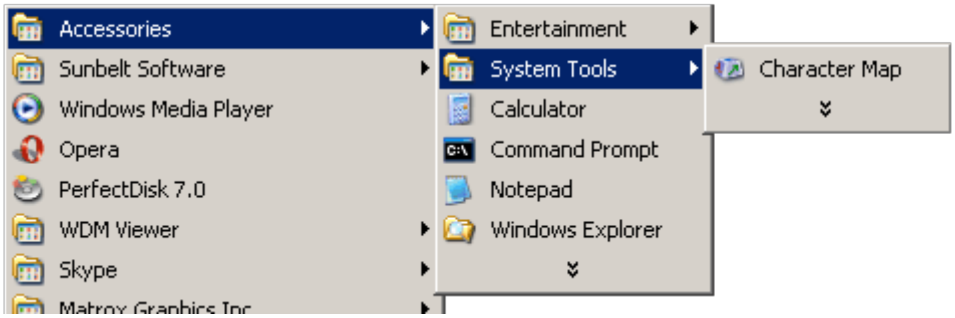
● Press the CAPS LOCK key ○ Press the SHIFT key

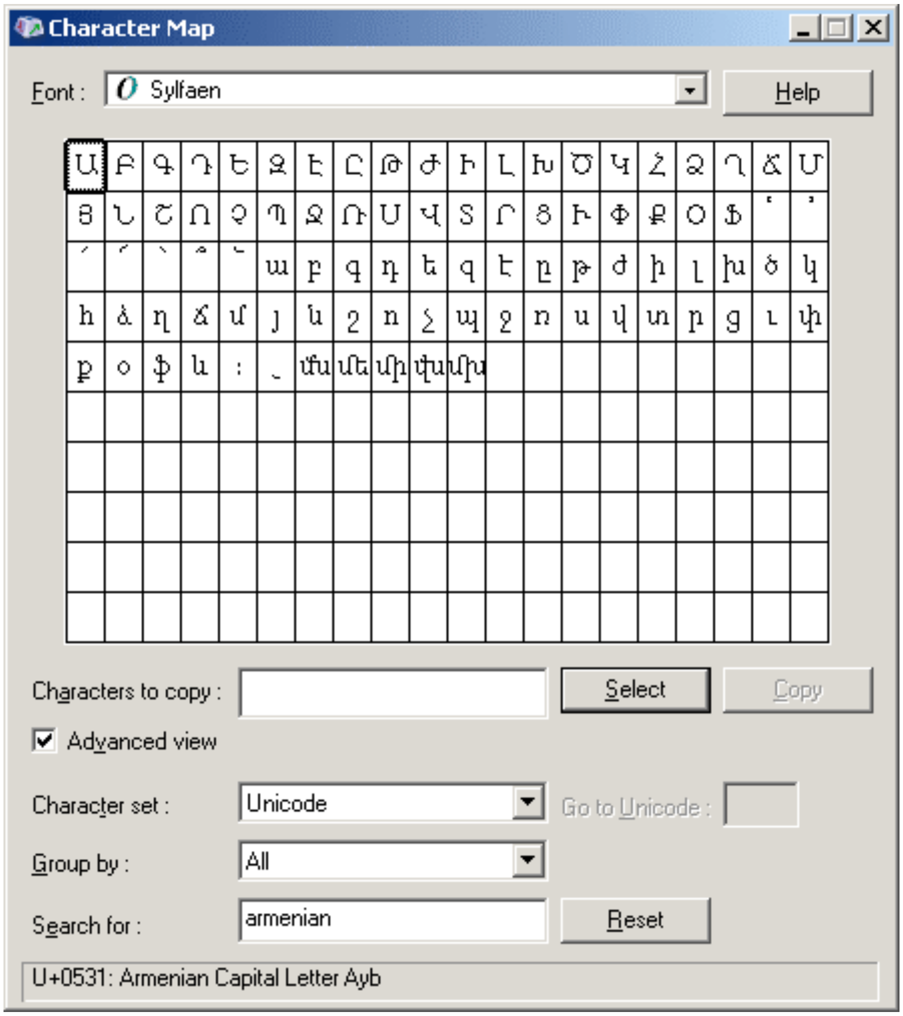
Hot keys for input languages

Action	Key sequence
Switch between input languages	Left Alt+Shift
Switch to Armenian - Armenian Eastern	(None)
Switch to Armenian - Armenian Western	(None)
Switch to English (United States) - US	(None)

4.4 Character Map

To access characters in the language that may not be available or easily accessible from the keyboard, open the Character Map (**Start > Accessories > System Tools > Character Map**).





Font : Sylfaen

Ա	Բ	Գ	Դ	Ե	Զ	Է	Ը	Թ	Ժ	Ի	Լ	Խ	Ծ	Կ	Հ	Ձ	Ղ	Ճ	Մ
Յ	Ն	Շ	Ո	Չ	Պ	Ջ	Ռ	Ս	Վ	Տ	Ր	Ց	Ւ	Փ	Ք	Օ	Ֆ	ՙ	՚
՛	՜	՝	՞	՟	ա	բ	գ	դ	ե	զ	է	ը	թ	ժ	ի	լ	խ	ծ	կ
հ	ձ	ղ	ճ	մ	յ	ն	շ	ո	չ	պ	ջ	ռ	ս	վ	տ	ր	ց	ւ	փ
ք	օ	ֆ	և	։	֊	ﬓ	ﬔ	ﬕ	ﬖ	ﬗ									

Characters to copy :

☑ Advanced view

Character set : Unicode

Go to Unicode :

Group by : All

Search for : armenian

U+0531: Armenian Capital Letter Ayb

Figure 11. The Windows XP Unicode Character Map Viewer

5 Further Reading

Website	Description
http://www.telf.com	Tools for converting older texts to Unicode, two very useful Unicode Armenian keyboard layouts, and a handful of Unicode fonts that include Armenian.
http://unicode.org/	Unicode Consortium website
http://en.wikipedia.org/wiki/Unicode	Wikipedia Unicode article
http://www.alanwood.net/unicode/	Alan Wood’s Unicode resources
http://www.wazu.jp/gallery/Fonts_Armenian.html	Unicode Armenian fonts
http://en.wikipedia.org/wiki/Armenian_alphabet	Armenian Alphabet
http://titus.uni-frankfurt.de/indexe.htm	TITUS: Thesaurus Indogermanischer Text- und Sprachmaterialien
http://www.deinde.org/unicode-for-mac/	Converting to Unicode
http://www.biblicalgreek.org/links/fonts/keyboard.html	Biblical Greek Site
http://www.stoa.org/unicode/index.html	Unicode Polytonic Greek
http://faculty.bbc.edu/rdecker/unicode.htm	Biblical Language Unicode
http://msdn.microsoft.com/en-us/library/ms776459(VS.85).aspx	Microsoft Unicode article
http://www.hayastan.com/fonts/	Windows Unicode setup
http://www.harkadir.am/index.jsp?sid=1&id=51&pid=1	Armenian keyboard layouts for Windows
http://www.microsoft.com/downloads/details.aspx?FamilyID=64FA945F-4204-4358-A172-F70B444A6F96&displaylang=en - Overview	Windows® XP հայերեն միջերեսի փաթեթ