The BU Phonetic keyboarding System

Albert Bickford, October 12, 2026
Updated October 12, 2026

1 Introduction

The BU[1] Phonetic Keyboard provides access to a wide range of characters for Latin-based scripts in Unicode (www.unicode.org), including:

- English, Spanish, French, German, and other major European languages[2]
- nearly complete set of IPA and Americanist phonetic symbols, include obscure and obsolete symbols
- special characters commonly-used in typesetting
- arrows
- common mathematical, numeric, and currency symbols

The BU Phonetic keyboard is one of the more extensive Unicode keyboards for Latin scripts available, although it still does not cover all of the hundreds of Latin characters in Unicode. I have tried to include those that are more likely to be used by linguists and others working with multiple languages.[3] To ease the memory load, the keyboarding conventions use a relatively small set of conventions that are applied very broadly and generally. Once you learn the conventions, you should be able to guess the keyboarding sequence for many characters without looking them up.[4] I have also tried to avoid using keystroke combinations that may be needed for other purposes, e.g. for shortcut commands in common application programs.

The BU Phonetic keyboard is not intended for everyone by any means. Most people have somewhat limited needs, for one particular language or a small group of them, or only for IPA. For them, this keyboard may be clumsy or unnatural to use and perhaps even be confusing or frustrating. Characters that they need to type often may require a series of several keystrokes. Such people may better be served by a smaller keyboard tailored specifically for their needs.[5] The BU Phonetic keyboard is intended for people who need to type a wide variety of Latin characters and other symbols, and who want to be able to do so with one single coherent system rather than constantly switching keyboards (either switching between multiple custom keyboards or turning a custom keyboard on and off). So, for example, it would be a good choice if you want to type practical orthographies for two or three languages plus phonetic transcription, as might happen when you are working on a text glossing or

---

[1] "BU" stands for "Bickford Unicode". I named it after myself not for vainglory but simply as an easy way to distinguish it from other Unicode keyboards. If you can think of a better name, I'm open to suggestions.

[2] I'm not sure that support for languages other than English and Spanish is complete yet, but will need help from someone who knows the needs of those languages to verify that all necessary characters are included. Note that "support" means only that the characters can be typed, not that you type them in the same way as you normally would in keyboards made specifically for those languages. Any keyboard that addresses the needs of multiple languages is going to have to deviate from the conventions normally used in standard keyboards for individual languages, especially if there is a desire to have one consistent system that applies to all languages equally.

[3] It was designed specifically for my own use as a linguist in the Mexico branch of SIL (https://mexico.sil.org/), but I'm releasing it more generally in hopes that it is useful to others. It is written so it can be easily modified by adding new characters or changing trigger keys, if necessary; see instructions and suggestions included in the source code (bu_phonetic.kmn).

[4] This, however, has the consequence of lengthening some of the key sequences from what might otherwise be possible; I opted for making it easier for the memory than for the fingers. In some cases shortcuts are available.

[5] For other Unicode keyboards, see www.keyman.com. If none of these other keyboards are adequate, you can try modifying this one, for example, by disabling key combinations that you don't need (see instructions inside the source code).

dictionary project. It will be especially helpful if, although you may need to type a wide variety of characters, you don't need to type any one character very often, as when preparing linguistic textbooks and other teaching materials. The goal has been to make it easy to *remember* how to type a large number of characters using a keyboarding system with a high degree of regularity and internal consistency, even if that means using more keystrokes for each character.

The BU Phonetic keyboard is based on the standard keyboard layout used in the United States and is designed to be used with "US" or "United States" as the underlying Windows keyboard.[6]

1.1 About this file

This file provides information on how to use this keyboard, as it is originally configured.[7] It also provides suggestions about how to use some of the Unicode characters that the keyboard generates.

This file provides complete, detailed, reference documentation. I have not attempted to provide a user's manual with reduced information for someone who doesn't need all the details, mostly because I don't know who is going to need it and what type of manual and selection of information they'll need. In the meantime, you may have to make your own: just make a separate list of the characters you generally use and how to type them, rather than trying to find them each time in this long list.

If you have a modified version of the keyboard, it should have been labelled as such, and the person who modified it for you should have provided you with a list of the modifications or a revised version of this file (in which case it too should be clearly identified as a revision of my original).

If you are viewing this file in PDF format, then all special characters should be displayed properly using fonts embedded in the PDF file.[8]

1.2 MIT License

© 2006-2026 J. Albert Bickford

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM,

[6] At this point, I do not know how it behaves when used with other underlying Windows keyboards. NEED TO TEST THIS.

[7] There are also some more esoteric things the keyboard does which are not documented here, and some things that are just plain cute but probably not useful. For information on these, see the source code for the keyboard.

[8] If you happen to have a version of this file in Microsoft Word format, you need Unicode fonts installed on your system that display all the characters. Doulos SIL is specified by this file for most of the characters. If you don't have this font and can't get it easily, then open "Tools, Options, Compatibility tab", choose "Font Substitution", and tell Word to substitute some other Unicode font that you do have, such as Arial Unicode MS. Some characters may not appear; for this reason, the PDF version of the documentation is preferable. However, the Word version has one advantage, at least in newer versions of Word: you can highlight a character, type ALT+X, and Word will change the character into the Unicode number that represents it. (Typing ALT+X a second time changes it back to the character.) That way, you can find out exactly what Unicode character is being generated by the table, without having to wade through the source code.

DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

1.3 Support

The most recent version is now available on the Keyman GitHub site at: https://github.com/keymanapp/keyboards/tree/master/release/b/bu_phonetic. You can write to me at albert_bickford@sil.org. I am willing to answer questions about it, fix bugs, and implement suggested improvements as time permits, but cannot promise any particular level of support. Existing keyboarding sequences may change from one version to the next, as more characters are added and things need to be rearranged to make room. I may not keep copies of all previous versions and do not intend to maintain a versioning system other than the file modification dates. I may not be able to provide much support for older versions, other than suggesting you switch to the most recent version (which may no longer have the same keyboarding conventions as the ones that you've gotten used to). But, I don't mind if you ask for any of these things.

2 How to use the BU Phonetic keyboard

2.1 Installation

If you are reading this file, you probably already have Keyman and the BU Phonetic keyboard installed on your system. If not, use the installer for the keyboard, which installs both it and Keyman.

2.2 Activating the keyboard

To activate the BU Phonetic keyboard, click on the Keyman icon in the system tray and choose it from the menu. When it is active, the BU icon will appear in the system tray.

The default hotkey to activate the keyboard is **CTRL+ALT+U**. This can be changed by right-clicking on the Keyman icon and choosing "Keyman configuration".

If you use the BU Phonetic keyboard with Microsoft Word, Word's AutoCorrect feature may make changes in addition to what this keyboard makes. Thus, the results in Word may not be as noted below, and this can be confusing. Consider turning off AutoCorrect if you use this keyboard with Word. (Depending on the version of Word, you do this in "Tools, AutoCorrect Options" or in "Tools, Options, Autocorrect tab".)

2.3 Trigger keys

By "trigger key", I mean a single primary keystroke (possibly in combination with SHIFT, CTRL, or ALT modifiers) that you type which transforms another character or is itself translated to something other than what it normally produces on a standard keyboard.

There are some general principles followed in this keyboard for the way trigger keys work. Knowing these principles can help you remember what the trigger keys mean and even enable you to guess how to key some characters without having to look them up.

- Trigger keys are typed after the letters they modify. This has several advantages compared to typing trigger keys before the character to be modified (which is what most people are used to):
  - It matches the order in which diacritics appear in Unicode (see section 5).
  - It provides better visual feedback when typing complex sequences of keys, because you can see the effect of each key immediately.
  - It makes it easier to implement a coherent and easily-remembered keyboarding system for a wide range of characters.

- To minimize interference with ordinary keyboarding, many of the trigger keys use the CTRL key as a modifier, and some also use SHIFT and/or ALT. In this file, this is indicated by notations like [CTRL+(]. In this notation, you do not use the SHIFT key unless it says specifically to do so. Thus, [CTRL+(] means to hold the CTRL key while typing the key that has (printed on it, *without* holding down the SHIFT key at the same time. In contrast, [CTRL+SHIFT+(] means to hold down both CTRL and SHIFT keys while typing the key that has (on it.
- The SHIFT key is used quite generally with trigger keys to mean "down". Thus, it is used to generate subscripts, place diacritics below base characters, and to invert characters. For example, [CTRL+SHIFT+|] adds a macron below the preceding letter (i.e. an underline: _), while [CTRL+|] puts it above the letter (¯).
- Many of the trigger keys involve rotation of a character in some way; rotation is always counterclockwise. For example, typing [CTRL+SHIFT+#] after an upwards arrow changes it into a northwest arrow. This principle also underlies the mnemonic value of many trigger keys for typing diacritics (see section 5).

2.4 Coping with unwanted changes

What happens when you want to type a series of characters that this keyboard translates to something else? I've tried to plan key sequences so this doesn't happen very often, but it will happen occasionally. There are several ways to cope. Three of them apply to any Keyman keyboard:

- You can turn off the keyboard. This is a bit clumsy if you need it turned off only for one character, but can be made easier by configuring Keyman so that the keyboard hotkeys toggle keyboard activation. Right-click on the Keyman icon in the system try, choose "Keyman Configuration", go to the Options tab, and put a check mark in front of "Keyboard hotkeys toggle keyboard activation". Then, you can quickly turn this keyboard off as well as on by typing CTRL+ALT+U (or whatever key sequence you have assigned to it).
- You can interrupt the keying sequence by typing something else that has no other effect. One thing that works well most of the time is to type the arrow keys [LEFT] and [RIGHT], one after the other. This has the effect of leaving the cursor where it was before, but if you are in the middle of a key sequence, it also interrupts the sequence so that the change does not take place. So, for example, this keyboard normally translates ae& to "æ". However, if you type ae[LEFT][RIGHT]&, you will get a literal "ae&", not "æ".
- If you *never* want to use a key sequence defined in the keyboard, you can have someone modify the keyboard to disable that sequence.

Some other techniques are unique to this keyboard:

- When the trigger key is an ordinary visible character (not a CTRL or ALT combination), you can usually undo the result by typing it a second time, so that you can get the literal sequence of characters without changes. For example, typing ae& normally produces "æ". But, if what you want in the file is a literal "ae&", you just type the ampersand a second time, and that will turn the "æ" into "ae&". Typing the ampersand a third time produces "æ&".[9]
- Other techniques for getting around certain specific problematic changes are listed in the charts below.

2.5 Other requirements for using Unicode

In order to use Unicode effectively, you need more than just a keyboard. Since Unicode may be new to you, this section summarizes what else you need in order to make use of this or any other Unicode keyboard.

The first thing you need is software that is designed to be used with Unicode. Usually, this means using fairly recent software (produced since about 2003). Here are some categories of Windows software and specific programs that work with Unicode:

[9] When the trigger key has CTRL or ALT modifiers, typing the trigger key a second time is normally ignored, although in some cases it may produce a double instance of the diacritic (e.g. double acute and double grave).

Category	Unicode-aware software
Operating system	Windows NT/XP[10]
Word processor	Microsoft Word 2003[11]
Desktop publication	Microsoft Publisher 2003[12]
	Adobe InDesign
	UltraXML
Scripture translation	Paratext 6.0
	SIL FieldWorks Scripture editor
Dictionary and text glossing	SIL Toolbox
	SIL FieldWorks Lexical and Text tools (currently under development)
Text processing utilities	SIL CC 8.0 (limited support)
	SIL TecKit and utilities that uses it
	XSLT processors

Most older software is not configured to work with Unicode. Such software can only access characters that are in a restricted character set called a "codepade", which usually has a maximum of about 220 characters available. The BU Phonetic keyboard will work with such software, but will only generate characters that are available in the standard codepage used with English versions of Windows.[13] This means that, in the charts of keystrokes in this documentation, the column labeled "Available characters" refers to the characters available in software that is configured to work with Unicode. Whenever a character is available in non-Unicode software, however, it is typed the same way as in Unicode software (with just a couple minor exceptions, noted below in the charts).

Next, you need one or more Unicode fonts. Several such fonts exist, and more are becoming available each year. Most Unicode fonts do not contain all the characters generated by the BU Phonetic keyboard, and some are more complete than others, so you need to test a font to make sure it has what you need.[14] Here are some fonts that include large numbers of Latin characters (not all of them include characters for the newest version of Unicode):

[10] Although some applications (such as Microsoft Office and Toolbox) can provide Unicode support on Windows 98/Me, most cannot, so Windows NT/XP is generally considered the minimum operating system that you need for good Unicode support.

[11] Word 97 had limited Unicode support, but it didn't start to be really useful until Word 2002, and positioning of diacritics was sometimes unsatisfactory until Word 2003, although technically this is because of the version of Uniscribe, not Word itself.

[12] I'm not sure how good the Unicode support is in Publisher 2002; Publisher 2003 is a safer bet.

[13] That is, codepage 1252, Windows Latin-1. It is possible that it may work with some other codepages.

[14] Some software, such as newer versions of Microsoft Word, is aware if a particular character is not available in the selected font, and automatically applies another font which does contain that character, using direct formatting. However, the font which Word selects may not be the most appropriate one, and it is better to develop a character style which you can apply in these cases. For example, in this file, there is a character style called "Unicode fallback" which uses the Code2000 font for those few characters that are not included in Doulos SIL and appears in this color.

Typeface name	Availability	Styles (bold, italic)[15]
(many of the standard Windows fonts, especially Arial Unicode MS and Lucida Sans Unicode)	(built-in)	depends on typeface
Doulos SIL	https://software.sil.org/doulos/	regular
Charis SIL	https://software.sil.org/charis/	regular, bold, italic, bold italic
Gentium	https://software.sil.org/gentium/	regular, italic
Everson Mono Unicode	http://www.evertype.com/emono	??

Most fonts made in the 1990s are not Unicode fonts; they only work with the characters in the system codepage.[16]

If any of the combinations of characters you use require special fine-tuning of their appearance (such as automatic positioning of diacritics over different heights and widths of base characters, or stacked diacritics), you will need a rendering engine that supports the special behavior you want, plus fonts and software that are enabled to work with it. On Windows, the standard rendering engine is called Uniscribe; for reliable positioning of diacritics that are encoded as separate characters (combining marks) you need the version of Uniscribe that was released with Office 2003, not the one that comes standard with Windows XP. Many, but not all, Unicode programs and fonts designed for recent versions of Windows will work adequately with Uniscribe. An alternative rendering engine is SIL's Graphite (scripts.sil.org), which provides specialized adjustments of appearance beyond what is available in Uniscribe. At the moment, it is only used in SIL's fonts and in the newest SIL software (WorldPad and FieldWorks), although efforts are underway to port it to other application software, particularly OpenOffice (a free alternative to Microsoft Office).

3 Phonetic and orthographic symbols

Most of the symbols in this section have uses in phonetic transcription,[17] or have been adapted from phonetic transciption for use in practical orthographies. There are, however, many other phonetic symbols in the next two sections. What distinguishes characters in this section is that they do not involve merely changes of size and position from other symbols (cf. section 4), nor do they involve diacritics (cf. section 5).

The following trigger keys are the ones most commonly-used in this section:

Type of symbol	Trigger character typed afterward
clicks	[CTRL+!]
ligatures	& (after a sequence of 2 or more characters)

[15] Windows is able to simulate bold, italic, and bold italic even when those styles are not provided by the font designer. However, the appearance is usually not as good as what a font designer can do, and there are some situations when the simulation is not available. For example, Adobe Acrobat can only display fonts in bold and italic if there is an actual font available in that style; it does not make use Windows' simulation of bold and italic.

[16] Some fonts, especially those used for minority languages, redefine the appearance of certain characters in a codepage as a way of providing a selection of characters different from what the codepage provides. These are called "hacked fonts" or "fonts with custom encodings". However, they are still limited to the same number of characters as in the codepage, and only the appearance of the characters change, not their behavior (upper- and lower-case pairs, word-building vs. punctuation, etc.).

[17] Upper-case versions of phonetic symbols are sometimes provided, even if the upper-case version only has orthographic (not technical) use.

Type of symbol	Trigger character typed afterward
reversed	[CTRL+#]
turned and inverted	[CTRL+SHIFT+#]
miscellaneous phonetic symbols derived from Greek (not in the above groups and not derived by application of diacritics)	[CTRL+/]
characters with palatal curls	[CTRL+SHIFT+/]
other phonetic symbols (not in the above groups)	[CTRL+&]

Usually, the mnemonic value of the key sequence is based on the typographic appearance of the special character, but in a couple cases it is based on the phonetic value of the character.

Some special characters are typed by modifying other special characters. So, for example, to get a reversed open e (epsilon) "ɜ", you first type the combination for an open e—e[CTRL+/]—then type [CTRL+#] to reverse it. In other words, e[CTRL+/][CTRL+#] produces "ɜ". You can often type the modifiers in more than one order; for example, e[CTRL+#][CTRL+/] also produces "ɜ".[18] To simplify the instructions below and make it clearer how the system works, many of these special characters are listed as themselves in the "What to type" column for other characters rather than repeating how to type them each place they appear. In these cases, you'll have to look at the main entry for typing that particular special character to find out how to type it. (When there is more than one sequence of trigger keys that can be used for the same character, it shows up in more than one place in the table.)

Sometimes there are shortcut ways of typing certain common characters, in addition to the longer and more general key sequence. Possibilities for modifying the keyboard to create additional shortcuts are noted in footnotes.

Description	Available characters	What to type
dotless exclamation mark[19] (="saltillo", enlarged straight apostrophe, orthographic glottal stop; cf. apostrophe, see below)	'	'[CTRL+&][20]
modifier letter apostrophes for ejectives (=glottalized consonants), sometimes also used for glottal stop in practical orthographies; cf. curved quotes, see below.	ʻ ʼ	'[CTRL+ALT+LEFT] '[CTRL+ALT+RIGHT] That is, type the sequence for a dotless exclamation mark, then type one of the directional modifiers that are also used with arrows. Be careful not to start with a straight apostrophe, or you will get a curved quote (see below), which looks almost the same but behaves like a punctuation character.

[18] Sometimes alternative orders are not available, if the required intermediate characters are not in Unicode. That is, each trigger key along the way must produce a valid Unicode character that is supported by the keyboard.

[19] In the most recent version of the keyboard, this produces U+A78C LATIN SMALL LETTER SALTILLO, which was added to the Unicode standard in April 2008 with Unicode 5.1. Older versions of the keyboard may produce other codepoints such as U+02C8 MODIFIER LETTER VERTICAL LINE or the SIL PUA character U+F21D DOTLESS EXCLAMATION.

[20] If desired, the keyboard can be easily modified so that typing apostrophe by itself will produce saltillo. In that case, '[BACKSPACE] or '' should produce an ordinary straight apostrophe. Search for "PATCH AREA saltillo" in the source code.

Description	Available characters	What to type
glottal stop (cf. question mark, see below)	ʔ	?[CTRL+&][21]
miscellaneous phonetic and other symbols based on Greek letters (other than those that fit in a more specific category, either in this section or the diacritic section)	ɑ ᵅ (“alpha”)	a[CTRL+/] ᵃ[CTRL+/]
	ɒ ᵖ (turned alpha)	ɐ[CTRL+/] ᵉ[CTRL+/]
	β ᵝ	b[CTRL+/] ᵇ[CTRL+/]
	ɛ ᵋ (open e)	e[CTRL+/] ᵉ[CTRL+/]
	Ɛ ɛ̨	E[CTRL+/] ɛ̨[CTRL+/]
	ɜ ᵊ (turned open e)	ə[CTRL+/] ᵊ[CTRL+/]
	ɜ ᵊ (reversed open e)	ɘ[CTRL+/]
	ɵ (closed open e)	e[CTRL+/][CTRL+/]
	ɞ (reversed cl. open e)	ɘ[CTRL+/][CTRL+/]
	Ɣ ɣ ˠ	G[CTRL+/] g[CTRL+/] ᵍ[CTRL+/]
	ɩ ɪ ᶦ	I[CTRL+/] i[CTRL+/] ᶦ[CTRL+/]
	µ	m[CTRL+/]
	ɸ ᵠ	p[CTRL+/] ᵖ[CTRL+/]
	Ʊ ʊ ᵘ	U[CTRL+/] u[CTRL+/] ᵘ[CTRL+/]
	ʉ	ʉ[CTRL+/]
	χ ˣ ₓ	x[CTRL+/] ˣ[CTRL+/] ₓ[CTRL+/]
letters with curls indicating sounds in the palatal region. (Don’t confuse these with the palatal hook diacritic, below.)	ɕ ᶜ ɖ ʥ[22] ʝ ʲ ɭ ɳ ʃ ʈ ʑ ᶻ ʒ	*letter* [CTRL+SHIFT+/]
miscellaneous phonetic symbols (other than those that fit in a more specific category)	ɬ ɱ ᵐ	l[CTRL+&] m[CTRL+&] ᵐ[CTRL+&]
	Ɲ ŋ ⁿ	N[CTRL+&] n[CTRL+&] ⁿ[CTRL+&]
	Ɲ ɲ ᶯ	Ŋ[CTRL+&] ŋ[CTRL+&] ᵑ[CTRL+&]
	ɾ	r[CTRL+&]
	Ʊ ʊ ᵛ	V[CTRL+&] v[CTRL+&] ᵛ[CTRL+&]
	Σ[23] ʃ ˢ	S[CTRL+&] s[CTRL+&] ˢ[CTRL+&]
	Ʒ ʒ ᶾ	Z[CTRL+&] z[CTRL+&] ᶻ[CTRL+&]
hooktop letters (implosives et. al.)	Ɓ ɓ Ƈ ƈ Ɗ ɗ Ɠ ɠ ɦ ʰ K ƙ Ƥ ƥ ƈ Ƭ ƭ Ƴ ƴ	*letter* [CTRL+&]
clicks	ʘ ǀ[24]	p[CTRL+!] s[CTRL+!]
	ǂ ǂ	t[CTRL+!] t[CTRL+SHIFT+#]
	ǁ	l[CTRL+!]
	ɔ	ʔ[CTRL+SHIFT+#]
	ƈ ![25]	c[CTRL+!] r[CTRL+!]

[21] If desired, the keyboard can be easily modified so that typing ? will produce glottal stop. In that case, typing ?[BACKSPACE] or ?? should produce an ordinary question mark. Search for “PATCH AREA glottal stop” in the source code.

[22] The digraph ʨ is also available in the keyboard, but not by this method of keyboarding, since a tc digraph does not exist in Unicode. The curl must be put on the c first, and then the t and ɕ can be combined in a digraph (see the section of the table that covers ligatures and digraphs).

[23] This is U+01A9 LATIN CAPITAL LETTER ESH, not U+03A3 GREEK CAPITAL LETTER SIGMA.

[24] The vertical bar used to represent a dental click is not the same Unicode character as the ordinary vertical bar (available on a normal English keyboard). The ordinary vertical bar is used, among other things, to represent a minor (foot) group in IPA or

Description	Available characters	What to type
ligatures and digraphs[26]	Æ æ ʣ DŽ Dž dž ʤ ʥ ʩ IJ ij LJ Lj lj ʪ ʫ ʨ NJ Nj nj Œ œ ʦ ʧ ʨ ʦ	*letter letter* &
	ɧ ƛ ɯ ɻ	ɧ& tl& uɥ& rl&
	ʜ ʜ	Hv& hv&
	Ŀ ŀ	L.[CTRL+^]& l.[CTRL+^]&[27]
	Ƣ ƣ	Oi& oi&
	Ȣ ȣ	Ou& ou&
	Þ þ	Th& th&
turned and inverted[28]	ɐ ᵊ ɒ ᵓ æ Ɔ ɔ ə[29] ᵊ ₔ ą ɐ ɜ ᵌ ɞ ɥ ¡ ᶤ Ш ɯ ᶣ ɹ ᵣ ʁ ᵡ ʁ ɻ ʌ ^ ʍ ʎ ʖ ʂ	*letter* [CTRL+SHIFT+#]
	Ə	E[CTRL+SHIFT+#]
	ʁ	R[CTRL+SHIFT+#] (shortcut)
	ɤ	ʁ[CTRL+#][30] (alternate method)
	ɻ	ɾ[CTRL+SHIFT+#]
reversed[31]	Ǝ ᵌ ɘ ɜ ᵌ ʒ ɷ и ᴎ я Ɛ ɕ ʕ ʕ ʿ	*letter* [CTRL+#]
	я	R[CTRL+#] (shortcut)
	ɿ	ɾ[CTRL+#]
	ʕ	?[CTRL+#] (shortcut)
dotless letters	ɪ ȷ	*letter* [CTRL+.]
tone letters	˩ ˨ ˧ ˦ ˥	*digit(1-5)*|&
	ꜜ ꜛ ˧ ˦ ˥	|*digit(1-5)*&
tone contours (these display as contour tone letters only in Doulos SIL;[32] other fonts show a sequence of level tone letters)	˩˩ ˩ ˨ ˧ ˦ ˩ ˧˩ ˨ ˧ ˦ ˩ ˨ ˧˨ ˧ ˦ ˩ ˨ ˧ ˧˥ ˥ ˩ ˨ ˧ ˦ ˥˥ ˨˩ ˧ ˦ ˥ etc.	sequence of 2–3 tone letters; once the first tone letter is typed, the 2nd and 3rd letter in the contour can be added just by typing the digit alone; it is not necessary to repeat the | and &.

---

the "pipe" character in some computer languages.

[25] The exclamation mark used to represent a retroflex click is not the same Unicode character as the ordinary exclamation mark. So, for example, typing r[CTRL+!][CTRL+SHIFT+#] does not produce ¡, like typing ![CTRL+SHIFT+#] does.

[26] This category includes a few characters that, technically-speaking, are not ligatures, but involve some sort of combination of typographic or phonological characteristics.

[27] In other words, L or l, followed by a middle dot (·), and then joined with &.

[28] "Turned" characters are those that are rotated 180°; "inverted" characters are those that are flipped vertically, like a reflection in a mirror held horizontally. As a mnemonic, think of the lines in the # character as containing horizontal and vertical mirrors.

[29] This is U+0259 LATIN SMALL LETTER SCHWA, not U+01DD LATIN SMALL LETTER TURNED E.

[30] In other words, to get ʁ (turned R), you start with a capital R, then either reverse and invert it, or invert and reverse it.

[31] "Reversed" characters are those that are flipped horizontally, like a reflection in a mirror held vertically. As a mnemonic, think of the lines in the # character as containing horizontal and vertical mirrors.

[32] As of this writing, tone contours that have the tone bar on the left do not display correctly in Word, although they do work in Notepad (when using the newest version of Uniscribe) and certain other applications.

Description	Available characters	What to type
IPA length and half-length marks	ː ˑ	:[CTRL+&] .[CTRL+&]
IPA marks for prosodic breaks and non-breaks	| ‖ ‿ ⁀	|[33] ||& ([CTRL+&])[CTRL+&]
arrows[34] (upstep, downstep, global rise and fall)	↑	|^[CTRL+ALT+UP]
	↓	|^[CTRL+ALT+DOWN]
	↗	|^[CTRL+ALT+RIGHT][CTRL+SHIFT+#]
	↘	|^[CTRL+ALT+DOWN][CTRL+SHIFT+#]

4 Size and position changes

Unicode provides a number of characters that are inherently superscripted or subscripted.[35] Many of them are called "modifier letters", which means that they are normally used as a supplementary symbol next to (usually after) some other character, such as ʰ for aspiration.[36] There is also a set of small capitals, and other miscellaneous characters that can be conceived as being larger or smaller versions of another character, without shifting it up or down.

Most of these modifications use the following trigger keys:

Type of symbol	**Trigger character typed afterward**
superscript version of character (or otherwise raised)	[CTRL+^]
subscript version of character (or otherwise lowered)	[CTRL+SHIFT+^]
larger version of character	[CTRL+ALT+^]
smaller version of character (e.g. small-caps)	[CTRL+ALT+SHIFT+^]

The modifier letter arrowheads are also included in this section; they are aimed in the desired direction using the same directional modifiers that are used for typing arrows (see section 8).

Description	**Available characters**	**What to type**
superscript letters[37]	ᴀ ᵃ ɐ ɑ ᴂ ᴃ ᵇ ᵝ ᶜ ᶝ ᴅ ᵈ ᶞ ᴇ ᴲ ᵉ ə ɛ ɜ ᶟ ᶠ ɢ ᵍ ɣ ʜ ʰ ɥ ʱ ɪ ᶦ ɨ ᶤ ɩ ᶥ ᴊ ʲ ɟ ʝ ᶨ ᴋ ᵏ ʟ ˡ ɭ ᶩ ɮ ᴍ ᵐ ɱ ᶬ ɯ ᶭ ɰ ᶰ ɴ ᶮ ᵰ ᶯ ɲ ᶯ ᵑ ᶲ ᴏ ᵒ ɵ ᶱ ɔ ᴘ ᵖ ᵩ ʀ ʳ ɹ ᶫ ɻ ˢ ᶳ ʃ ᵀ ᵗ ʇ ᶿ ᵾ ᵘ ᶸ ᵫ ᵿ ᶶ ʊ ᵛ ᶹ ʌ ᵂ ʷ ˣ ᵡ ʸ ᶻ ᶼ ᶽ ᶾ ʒ ʔ ˤ ʻ	*letter* [CTRL+^]

[33] The single vertical bar is accessible from an ordinary English keyboard, but is included here for completeness, to help those who aren't sure which Unicode character to use to represent this IPA character.

[34] These are the ones that have use for phonetics; see section 8 for the full set.

[35] To be precise, not all of these characters are superscripts and subscripts; a few are more vaguely conceived as being raised or lowered versions of other characters.

[36] Most of the characters designated as modifier letters are superscripts, although a few are subscripts. Whether a character is designated as a modifier letter does not seem to make much practical difference in terms of its actual behavior; it seems rather to just be a useful naming convention for distinguishing superscripted characters from their full-size counterparts.

[37] a) Of the three superscript open e's, the first is regular, the second is turned (rotated 180°), and the third is reversed (flipped right-to-left). b) In some cases there is a superscript small capital letter in addition to the superscript capital, e.g. for I, L, and N. c) The superscript reversed glottal stop is U+02C1 MODIFIER LETTER REVERSED GLOTTAL STOP, not U+02E4 MODIFIER LETTER SMALL REVERSED GLOTTAL STOP.

Description	Available characters	What to type
	ʔ	ɬ[CTRL+^] (shortcut)
		?[CTRL+^] (shortcut)
Spanish feminine sign (superscript underlined a)	a̲	a̲[CTRL+^] (Unicode)
		a[CTRL+^] (non-Unicode)
Spanish masculine sign (superscript underlined o)	o̲	o̲[CTRL+^] (Unicode)
		o[CTRL+^] (non-Unicode)
superscript digits and other mathematical symbols	⁰ ¹ ² ³ ⁴ ⁵ ⁶ ⁷ ⁸ ⁹	*digit* [CTRL+^]
	⁺ ⁻ ⁼ ⁽ ⁾	*symbol* [CTRL+^]
middle dot (Americanist half-long)[38]	·	·[CTRL+^]
other superscript symbols	ː	ː[CTRL+^]
modifier letter vertical lines (primary and secondary stress)	ˈ	ˈ[CTRL+^]
	ˌ	ˌ[CTRL+SHIFT+^]
subscript letters	ₐ ᵦ ₑ ₔ ᵢ ₒ ᵣ ᵤ ᵥ ₓ ᵪ	*letter* [CTRL+SHIFT+^]
	ᵞ	ᵞ[CTRL+SHIFT+^]
	ᵩ	ᵩ[CTRL+SHIFT+^]
subscript digits and other mathematical symbols	₀ ₁ ₂ ₃ ₄ ₅ ₆ ₇ ₈ ₉	*digit* [CTRL+SHIFT+^]
	₊ ₋ ₌ ₍ ₎	*symbol* [CTRL+SHIFT+^]
letters made larger (i.e. upper case versions of special characters)	Æ Ǝ Ə Ɛ Ɣ Ɯ Ŋ Ɲ Ɔ Œ	*letter* [CTRL+ALT+^]
	ɩ (iota)	ɩ[CTRL+ALT+^]
letters made smaller (most of these are small-caps, based on capital letters)	ʙ ɢ ɠ ʜ ɪ ˈ ɬ ʟ ˡ ɭ ɴ ᴺ ʀ ᴜ ᵘ ʏ ɶ ȣ	*letter* [CTRL+ALT+SHIFT+^]
	ɣ	ɣ[CTRL+ALT+SHIFT+^]
modifier letter arrowheads (Americanist fronting, backing, raising, lowering)	<	^[CTRL+ALT+LEFT]
	>	^[CTRL+ALT+RIGHT]
	ʌ	^[CTRL+ALT+UP]
	ʏ	^[CTRL+ALT+DOWN]

Caution: These size and position variations should not be used interchangeably with superscript, subscript and small-cap formatting. You should only use a size or position variant of a character if it should *always* appear that way, regardless of formatting, or if its appearance as a superscript/subscript/small-cap cannot be predicted by the computer based on other factors. So, for example, these characters are appropriate for the following uses:

- phonetic modifications of sounds (aspiration, palatalization, labialization, etc.)
- tone numbers (since you want them to stay as superscripts even if formatting is stripped off, e.g. when Microsoft Word automatically puts them in page headers in a dictionary)
- homophone numbers in a formatted dictionary (same reason, but see below for dictionary databases)
- small-caps used as phonetic symbols

In many cases, superscripting and subscripting should be done by formatting rather than by using modifier letters. This includes the following situations:

[38] The Unicode character U+00B7 MIDDLE DOT is a punctuation character, and as such may not be the best choice to represent the half-length mark in the Americanist tradition. A better choice is U+02D1 MODIFIER LETTER HALF TRIANGULAR COLON, which is used in the IPA to indicate half-length and sometimes also in the Americanist tradition. Some fonts may even provide a round dot as an alternate glyph for the half triangular colon. (At the time of this writing, relatively few word processing and publication programs can access the alternate glyphs in a font, but hopefully alternate glyphs will be more accessible in the near future.)

- footnote numbers (since they are sometimes formatted as full-size numbers inside brackets; plus, even when they are superscripted, the superscripting is generated automatically by word processors using formatting)
- homophone numbers in a dictionary database (since the superscripting can be applied automatically to the number as formatting, either within the database itself or when the database is converted to a formatted dictionary)
- mathematical formulas (since superscript and subscript formatting can be applied automatically by an equation editor)
- small-caps used on morpheme glosses (e.g. in a glossed text)

If you're not sure whether to use superscript/subscript/small-cap characters or formatting, get advice from someone who knows electronic publication well.

5 Diacritics

Unicode provides two ways of representing diacritics:

- composite characters: a single character that represents a combination of a base character plus one or more diacritics
- combining marks: a diacritic that is represented as a separate character; the computer automatically positions the diacritic above or below the base character which precedes it.

There are many composite characters in Unicode, but by no means does it include all possible combinations of base characters and diacritics. Many combinations have to be represented with combining marks, and some diacritics are available only as combining marks.

The BU Phonetic keyboard generates both composite characters and combining marks, with preference given to composites whenever they are available.[39] With newer software (such as Word 2003[40]) that knows how to handle combining marks intelligently (making them behave for the most part as if they part of composites), you probably will not notice the difference most of the time. However, there are times when it is helpful to be aware of the difference (e.g., when using `[BACKSPACE]` to remove a diacritic, or when typing a diacritic over a no-break space to get a diacritic in isolation). At these times, you need to remember that diacritics are always placed *after* the letter they apply to. That's one of the main reasons diacritics are typed after the base character in this keyboard.[41]

There are several other conventions for how trigger keys are used to generate diacritics:

- In most cases, diacritics that occur above the base character are typed using the `CTRL` key in combination with some other key; those that appear below the base character are typed using `CTRL+SHIFT`. (Reminder: in combination with `CTRL`, only press the `SHIFT` key if the table below specifically says to do so, even if the main key that is listed is normally typed with the `SHIFT` key.)
- A few diacritics require you to type `ALT` or `CTRL+ALT` in combination with some other key, but the same convention is maintained—trigger keys with `SHIFT` produce diacritics below the letter.
- Some diacritics require you to type `%` before the trigger key; these are the more unusual ones.
- A few trigger keys, especially those that are normally visible, generate diacritics only with certain base characters. The base characters that accept each diacritic are listed in the table below, in the column labeled

[39] There are a few obscure composite accented letters in Unicode, mostly in the ranges 0100–01FF (Latin Extended-A/B) and 1E00–1EFF (Latin Extended Additional), which are not currently generated by this table, but these combinations can still be achieved using combining marks as long as the base characters and diacritics are available separately.

[40] In Word, see "Tools, Options, Complex Scripts tab". On this tab, you can control whether combining marks are displayed at all and whether to display them with a distinct color.

[41] For a fuller discussion of the reasons, see section 2.3.

"Available combinations". With these diacritics, typing % before the trigger key will force the diacritic to be inserted,[42] even on characters that don't normally accept it.[43] If there are no specific combinations listed under "Available combinations", then the diacritic is available on all characters regardless of how it is typed.[44]

- In many cases, there's a simple way to remember which trigger keys represent which diacritics. You can think of giving the trigger key a quarter turn counterclockwise and then positioning it above, below, or through the base character. Thus,

¼ turn counterclockwise on a trigger key that includes the following:	produces these diacritics:			
:	¨	◌̈		(umlaut)
|	¯	◌̄	-	(macron)
(	˘	◌̆		(breve)
)	̆	◌̑		(inverted breve)
[		◌̪		(inverted bridge)
]	̪	◌̺		(bridge)
{	ˇ	◌̌		(caron/wedge)[45]
}	̂	◌̂		(circumflex)

[42] Some of the diacritics that can be forced into the data using % are inserted as combining marks which are not recommended for normal use (because of unsatisfactory appearance in most fonts). These are identified in the tables as being available on "any letter, but not recommended". It is now generally regarded that it was a mistake to have included combining marks in Unicode for overlay diacritics, because it is very difficult for font designers to position them correctly on the full range of base characters that they could conceivably be used with. As such, the composite characters in Unicode that use these diacritics are not considered to be equivalent to sequences that use combining marks. The same is true for the hooks that attach to base characters (palatalization, retroflexion, rhoticization). These diacritics should normally only be used when they are available in composite characters, such as those noted in the table under "Available characters". The Unicode Technical Committee (UTC) is still accepting composite characters into Unicode if they contain overlay diacritics (as long as they can be shown to have a history of actual use) and hooks, even though there is a general policy against introducing new composite characters. Doulos SIL includes a number of composite characters in the SIL Private Use Area (PUA) which have already been accepted for future versions of Unicode; their PUA codepoints could be easily patched into this keyboard if necessary. (I have not done so, however, since I'm waiting for their official inclusion in the standard before supporting them.) I have included the ability to keyboard these diacritics as combining marks only so there is a way to type existing published data that may contain base+diacritic combinations which are not yet available in Unicode or in the PUA as composite characters. If you discover any such combinations in actual use, please notify the SIL PUA committee via http://scripts.sil.org, or send me an email (albert_bickford@sil.org) and I will do so. However, you should avoid using these problematic combining marks in novel combinations with base characters (i.e., combinations with no historical precedent), since it is less likely that the UTC would look favorably on such proposals.

[43] Typing % in front of a trigger key can also be used to stack a second instance of a diacritic in combination with an existing one, if for some strange reason that is ever necessary. It also makes it possible to type a diacritic on a no-break space, which is not normally possible with single diacritics (but see footnote 48). In general, when a diacritic is normally typed with a single trigger key, you can also type it by preceding the trigger key with %. The effect is the same, except that certain restrictions are removed when using the longer sequence with %.

[44] In such cases, some of the generated characters may be composites whereas others will use combining marks; this distinction is not noted in the tables, as long as the combining marks generate good results. However, with overstrike diacritics and certain hooks, the combining marks do not generate good results. In such cases, all the available composite characters are listed in the "Available combinations" column and although the combining mark can be generated by the keyboard on other base characters, this is not recommended for normal use. If you need such a combination on a regular basis, you should consider requesting that it be added to Unicode.

[45] For the curly brackets { } (braces), what counts is the point in the center.

¼ turn counterclockwise on a trigger key that includes the following:	produces these diacritics:
=	" " (double vertical line)

Further rationale for the choice of certain trigger keys is provided in footnotes below.

- When the trigger key is an ordinary printable character (not a CTRL or ALT combination), typing the trigger key two or three times allows you to type all possible combinations of ordinary/accented letter plus the diacritic key. Thus:

Combination	What to type
ã	a`
a`	a``
ã`	a```

When the trigger key is a CTRL and/or ALT combination, typing the trigger key a second time is normally ignored and there is a beep.

- If you type [BACKSPACE] immediately after the diacritic is typed, the diacritic will be removed, even with composite characters.[46]

Diacritics can be typed on a wide variety of base characters. Three special situations should be noted:

- To get diacritics on special characters, type the key sequence for the special character first, followed by the trigger key for the diacritic. Thus, to get æ̀, you first type ae& to produce "æ", then type [CTRL+`] to add the grave accent. Once you start adding diacritics, you cannot change the base character.
- To put multiple diacritics on the same base character, type the ones that appear above the character first, going upward, then the ones that appear below the character, going downward.[47] Thus, to type ạ̋,
  - type a,
  - then [CTRL+:] for the diaeresis,
  - then [CTRL+'] for the acute,
  - then [CTRL+SHIFT+|] for the macron below,
  - and finally [CTRL+SHIFT+.] for the dot below.
- To type a diacritic in apparent isolation (not on a base character), precede it by a no-break space ([CTRL+SHIFT+SPACE]) and %.[48] For example, [CTRL+SHIFT+SPACE]%[CTRL+'] produces an acute accent ´ by itself. The keyboard tries to prevent you from typing diacritics on ordinary spaces, as this may result in a

[46] Once you've moved the insertion some place else and come back to a character, then this special function of [BACKSPACE] is not available, as with all the special key sequences in a Keyman keyboard. However, you will still be able to use [BACKSPACE] to erase the diacritic from the base character if the diacritic is a combining mark. On a composite character, [BACKSPACE] will erase the whole combination, base character plus diacritic. In programs like newer versions of Word that make combining marks behave in many ways like composites, [BACKSPACE] will always remove the last diacritic that is encoded as a combining mark; the only way to remove intermediate diacritics is to delete the whole combination and retype it. In older programs that allow you to position the insertion point between the base character and a combining mark, you may be able to do fine surgery by deleting any of the combining marks individually (or even deleting the base character, which will move the combining mark(s) to the previous character), but it can often be hard to tell exactly where the insertion point is located. It is probably easiest just to delete the whole combination and retype it.

[47] Unicode specifies that all diacritics above a character should be encoded before all diacritics below, and that if there is more than one diacritic above or below, they are encoded in order from the base character outward. If you type them in a different order, you may still get a result that looks right, but it may not be handled properly by some programs.

[48] If you need to type diacritics in isolation frequently, the keyboard can be easily modified so that you don't need to type % in order to put a diacritic on a no-break space. For instructions on how to do so, look in the source code for the keyboard and search for "PATCH AREA diacritics in isolation".

line break between the space and the diacritic, which will leave the diacritic hanging out beyond the left margin on the second line.[49] There are ways to fool the keyboard into letting you do this anyway, but I recommend that you don't try to sneak diacritics in where they don't belong: after ordinary spaces, at the beginnings of lines, on punctuation marks, etc.

5.1 Diacritics above the base character

Description (and typical use)	Available combinations	What to type
acute (stress, high tone)	x́ (any letter)	*letter* [CTRL+']
grave (low tone)	x̀ (any letter)	*letter* [CTRL+`]
tilde[50] (nasalization)	Ãã Ẽẽ Ĩĩ Ññ Õõ Ũũ Ỹỹ	*letter* `[51]
	x̃ (any letter)	*letter* %`
diaeresis (umlaut)	ẍ (any letter)	*letter* [CTRL+:]
macron (long, mid-tone)	x̄ (any letter)	*letter* [CTRL+|]
double macron	x͞x (any pair of letters)	*letter* %|*letter*
caron (=hacek, wedge; low-high tone, Americanist alveopalatals)[52]	x̌ (any letter)	*letter* [CTRL+{]
circumflex (high-low tone)	x̂ (any letter)	*letter* [CTRL+}]
double acute (extra high tone)	x̋ (any letter)	*letter* [CTRL+'][CTRL+']
double grave (extra low tone)	x̏ (any letter	*letter* [CTRL+`][CTRL+`]
breve (short)	x̆ (any letter)	*letter* [CTRL+(]
inverted breve	x̑ (any letter)	*letter* [CTRL+)]
double breve	x͡x (any pair of letters)	*letter* %(*letter*
double inverted breve (=ligature tie above: double articulations, affricates)	x͜x (any pair of letters)	*letter* %)*letter*
bridge above	x̄ (any letter)	*letter* [CTRL+ALT+]]
ring above	x̊ (any letter)	*letter* [CTRL+*]
comma above (Greek smooth breathing)	x̓ (any letter)	*letter* %c
reversed comma above (Greek rough breathing)	x̔ (any letter)	*letter* %c[CTRL+#]
turned comma above (also called "cedilla above")[53]	x̒ (any letter)	*letter* [CTRL+,] or *letter* %c[CTRL+SHIFT+#]
vertical line above	x̍ (any letter)	*letter* %'

[49] As of Unicode 4.1.0 the use of combining marks on ordinary spaces is no longer recommended, even though the spacing diacritics in Latin-1 and some of the modifier letters still have compatibility decompositions to space plus combining mark.

[50] On many keyboards, the backquote ` (usually used to type a grave accent) is on the same key as the tilde. Since the tilde is fairly common, especially in Spanish, I opted to make it easier to type than the grave accent. Thus, I used backquote to type tilde and [CTRL+`] for the grave accent. These choices can be changed in the table easily if you prefer something else.

[51] Or, to put it another way, type the key that has tilde on it, but *don't* use the shift key with it, following the general convention that diacritics above do not use SHIFT and diacritics below, do.

[52] With d, L, l, and t, a caron may look like an apostrophe rather than a wedge, depending on the font.

[53] When a composite character with a turned comma above has "WITH CEDILLA" in its name, it may also be generated with the key for cedilla (listed among the diacritics below).

Description (and typical use)	Available combinations	What to type
double vertical line above	x̎ (any letter)	*letter* [CTRL+=]
left angle above (unreleased)	x̒ (any letter)	*letter* %[CTRL+#]
dot above (former IPA for palatals)	ẋ (any letter)	*letter* [CTRL+.]
	ı ȷ	i[CTRL+.] j[CTRL+.][54]
x above (mid-centralized)	x̆ (any letter)	*letter* %x

5.2 Diacritics through or next to the base character

Description (and typical use)	Available combinations	What to type
stroke (=bar through)	ʙ ᵝ ƀ Đ đ ᵟ Ꞡ ꞡ Ħ ħ ɨ ɨ ᵻ ⁱ ɨ Ɬ ɬ Ɵ ɵ ᵠ ɸ ʃ Ŧ ŧ ʉ ᵿ ᵾ ᵿ Ƶ ƶ ʕ ʔ ʖ	*letter* [CTRL+ALT+|]
	ɟ	j[CTRL+ALT+|]
	θ ᶿ	0[CTRL+ALT+|] ᵒ[CTRL+ALT+|] ("zero with stroke")
	ʙ	B[CTRL+ALT+|] (shortcut)
	ᵿ	U[CTRL+ALT+|] (shortcut)
	ʔ	?[CTRL+ALT+|] (shortcut)
	x̶ (any letter, but not recommended[55])	*letter* %[CTRL+ALT+|]
tilde overlay	ƀ đ ɬ ꝑ ꝓ ᵽ ɾ ɾ ʂ ŧ ƶ	*letter* [CTRL+ALT+~][57]
	x̴ (any letter, but not recommended[56])	*letter* %[CTRL+ALT+~]
solidus (=slash)	Ⱥ ȼ ɇ Ł ł Ø ø Ɉ	*letter* [CTRL+ALT+/]
	X̸ (any letter, but not recommended[58])	*letter* %[CTRL+ALT+/]
rhotic hook	ɚ ɝ	ə%r ɜ%r
	x˞ (any letter, but not recommended[59])	*letter* %r

5.3 Diacritics below the base character

Description (and typical use)	Available combinations	What to type
acute below	x̗ (any letter)	*letter* [CTRL+SHIFT+']
grave below	x̖ (any letter)	*letter* [CTRL+SHIFT+`]

[54] In other words, using the key for the dot diacritic on i or j has the effect of *removing* the dot, making a dotless i (e.g. for Turkish) or dotless j. If for some reason you want a second dot on an i or j, you must force it with the % key. However, your software (like some of mine) may place the second dot in the same place as the first, making it look like there is just one dot; this is presumably a bug.

[55] See footnote 42.

[56] See footnote 42.

[57] Remember, when there are CTRL or ALT modifiers listed, don't type the SHIFT key unless it says so explicitly.

[58] See footnote 42.

[59] See footnote 42.

Description (and typical use)	Available combinations	What to type
tilde below (creaky voice)	A̰ a̰ Ḛ ḛ Ḭ ḭ O̰ o̰ Ṵ ṵ Æ̰ æ̰ ɨ̰ ʉ̰	*letter* ~[60]
	x̰ (any letter)	*letter* %~
diaeresis below (breathy)	ẍ (any letter)	*letter* [CTRL+SHIFT+:]
macron below (=underline, line below: low tone)	x̱ (any letter)	*letter* [CTRL+SHIFT+|]
double macron below	x̱x̱ (any pair of letters)	*letter* %|[CTRL+SHIFT+^]*letter* (i.e., a double macron that has been lowered)
circumflex below (Americanist fronted)	x̭ (any letter)	*letter* [CTRL+SHIFT+}]
caron below	x̬ (any letter)	*letter* [CTRL+SHIFT+{]
inverted breve below (non-syllabic)	x̯ (any letter)	*letter* [CTRL+SHIFT+)]
breve below	x̮ (any letter)	*letter* [CTRL+SHIFT+(]
double breve below	x͜x (any pair of letters)	*letter* %([CTRL+SHIFT+^]*letter* (i.e., a double breve that has been lowered)
bridge below (dental)	x̪ (any letter)	*letter* [CTRL+ALT+SHIFT+]]
inverted bridge below (apical)	x̺ (any letter)	*letter* [CTRL+ALT+SHIFT+[]
square below (laminal)	x̻ (any letter)	*letter* ˽ or *letter* ˼
seagull below	x̼ (any letter)	*letter* %}
inverted double arch below (labialized)	x̫ (any letter)	*letter* %{
palatal hook (don't confuse this with curls for sounds in the palatal region, see above)	ɓ ɗ ƒ ɠ ƙ ɭ ᶩ ɱ ɳ ƥ ɾ ʂ ʃ ʈ ᶵ ʏ ɤ ʒ	*letter* %,
	ʒ̡ (any letter, but not recommended[61])	
retroflex hook	ɖ ɗ ɭ ᶩ ɳ ᶯ ɾ ɻ ᶦ ʂ ᶳ ʃ ʈ ʈ ʐ ᶻ	*letter* %;
	ɑ ɒ ɐ ɜ ɞ ɜ ɻ ɥ	
	ʒ̢ (any letter, but not recommended[62])	
cedilla (with some letters it appears as a comma below or even a turned comma above)[63]	x̧ (any letter)	*letter* [CTRL+SHIFT+,]

[60] Or, to put it another way, on an English keyboard, the shifted version of `, the key that generates a tilde above.

[61] See footnote 42.

[62] See footnote 42.

[63] The cedilla and the comma below are often used interchangeably in practical orthographies. As a result, although they do have distinct combining marks in Unicode, they may often both look a comma. With some base characters, a cedilla combining mark looks like either a cedilla or a comma below (depending on the font). Often in these cases, there is only a single composite character corresponding to the two diacritics, and it is called "WITH CEDILLA" but its glyph may look like either diacritic. On other base characters, the cedilla always looks like a cedilla, and there may even be a composite character for each diacritic (e.g. S and s). In one case, with T and t, there are separate composite characters with cedilla and with comma below, yet even then the composites that are named "WITH CEDILLA" may look like they have a comma below (e.g. in Times New Roman). (Finally, in at least one case, on g, the cedilla is represented as a turned comma above.) This means you have to be careful to key these diacritics correctly and consistently in order to get the right character with the right codepoint. When in doubt that you're getting the correct character, use Word's ALT+X command to examine the actual codepoint that is being generated, and look up the description of that character in Unicode documentation.

Description (and typical use)	Available combinations	What to type
ogonek (=polish hook, Americanist nasalization)	x̨ (any letter)	*letter* %n
ring below (voiceless)	x̥ (any letter)	*letter* [CTRL+SHIFT+*]
half-rings below	x̜ (any letter) x̝ (any letter)	*letter* [CTRL+SHIFT+*] [CTRL+ALT+LEFT] *letter* [CTRL+SHIFT+*] [CTRL+ALT+RIGHT]
dot below (Americanist backed)	ẋ (any letter)	*letter* [CTRL+SHIFT+.]
comma below[64]	x̦ (any letter)	*letter* %[SHIFT+c][65]
vertical line below (syllabic)	x̩ (any letter)	*letter* %[SHIFT+']
double vertical line below (strong articulation, fortis)	x̳ (any letter)	*letter* [CTRL+SHIFT+=]
left angle below (weak articulation, lenis)	x̗ (any letter)	*letter* %[CTRL+SHIFT+#]
snake below (Americanist weak articulation, lenis)[66]	x̠ (any letter)	*letter* %s
plus below (advanced)	x̟ (any letter)	*letter* %+
minus below (retracted)	x̱ (any letter)	*letter* %-
tacks below:		
left (advanced tongue root)	x̘ (any letter)	*letter* +[CTRL+ALT+←][67]
right (retracted tongue root)	x̙ (any letter)	*letter* +[CTRL+ALT+→]
up (raised)	x̝ (any letter)	*letter* +[CTRL+ALT+↑]
down (lowered)	x̞ (any letter)	*letter* +[CTRL+ALT+↓]
x below	x̽ (any letter)	*letter* %[SHIFT+x]

6 Punctuation marks and other special characters used in typesetting

The typing conventions for punctuation marks and other symbols used in high-quality typesetting are somewhat more varied and less systematic than other characters in the keyboard, but hopefully not difficult to remember. Some of them may be modified easily, as mentioned in the footnotes.

Description	Available characters	What to type
straight apostrophe (cf. dotless exclamation mark, see above)	'	'[68]
curved single quote marks & apostrophe (cf. modifier letter	‘ ’	'[CTRL+ALT+LEFT] '[CTRL+ALT+RIGHT]

[64] See footnote 63.

[65] In other words, %C generates a comma below, whereas %c generates a comma above.

[66] U+1DC2 COMBINING SNAKE BELOW does not display correctly even in Word 2003; the snake appears to the right of the letter it modifies, instead of underneath.

[67] These characters use the some directional modifier keys as the arrows and other graphical elements described in section 8.

[68] If the keyboard has been patched to use apostrophe by itself for typing saltillo (see footnote 20), then '[BACKSPACE] or '' should produce an ordinary straight apostrophe.

Description	Available characters	What to type
apostrophes for ejectives, see above)		That is, first type the sequence for a straight apostrophe, then type one of the directional modifiers that are also used with arrows. Be careful *not* to start with a dotless exclamation mark (saltillo, see above), or you will get a modifier letter apostrophe (for ejectives), which looks almost the same but won't behave like a punctuation character.
curved double quote marks	“	"[CTRL+ALT+LEFT]
	”	"[CTRL+ALT+RIGHT]
“low 9” quote marks	‚	'[CTRL+ALT+DOWN]
	„	"[CTRL+ALT+DOWN]
single angle quotation marks (=single guillemets)	‹ ›	<& >&[69]
double angle quotation marks (=chevrons, guillemets)	« »	<<& >>&
Question mark (cf. glottal stop, see above)	?	?[70]
Inverted ? and !	¿ ¡	*symbol* [CTRL+SHIFT+#]
Exclamation and question marks in combination	‼ ⁇ ⁈ ⁉	*symbol symbol* &
	⸘	?!&& or !?&&
hyphen minus (The traditional ASCII character that is often used ambiguously for either a hyphen or a minus sign.)	-	- (i.e. type it normally)
soft hyphen U+00AD (A hyphen used to mark an optional hyphenation point; only visible when at the end of a line. NB: *not* used in Microsoft Word for the optional hyphen.[71])	*not visible (nor effective) in Word*	-s&
hyphen (An unambiguous representation for a hyphen.[72] Not often used; usually the hyphen minus shown above is used instead.)	-	-h&
non-breaking hyphen (Looks exactly like a hyphen but will not allow a line-break. NB: *not* used in	-	-nb&

[69] If you need to type a literal >&, e.g. in UNIX-based programming languages for a shell redirection, type >&&.

[70] If the keyboard has been patched to use ? by itself for typing glottal stop (see footnote 21), then ?[BACKSPACE] or ?? should produce an ordinary question mark. This can then be optionally inverted with [CTRL+SHIFT+#].

[71] Microsoft Word uses U+001F SI for an optional hyphen.

[72] For an unambiguous minus sign, see section 7.

Description	Available characters	What to type
Microsoft Word for the non-breaking hyphen[73])		
figure dash (A dash used for separating two digits, since it is, in theory at least, the exact width of a digit.)	–	-f&
en dash (A dash that is one “en” in length. Often used in place of a figure dash, since it is typically about the same length or only slightly longer.)	–	==[74] or ==& or -n&
two equals signs	==	=== or ==[BACKSPACE]
line of equals signs	========	three or more equals signs
em dash (A dash that is one “em” in length. Generally used as an alternative to a semicolon or in pairs for parenthetical comments; cf. quotation dash below.)	—	--[75] or --& or -m&
two hyphens	--	--- or --[BACKSPACE]
line of hyphens	----------	three or more hyphens
quotation dash (A dash that is used for introducing a quotation. Typically about the same length as an em dash or slightly longer, but does not allow line breaking between it and a following word.)	—	-q&
ellipsis	…	...[76] or ...&
three periods	...	...[BACKSPACE][77]
line of periods	.....	five or more periods
daggers	† ‡	tt& TT&
section and paragraph signs	§ ¶	SS& PP&
miscellanous symbols	¦ ¬	|[CTRL+&] -[CTRL+&]
	‖ ‗	||[CTRL+&] __[CTRL+&]

[73] Microsoft Word uses U+001E RS for a non-breaking hyphen.

[74] If this behavior is undesirable, it can be easily suppressed; see “PATCH AREA en dash” twice in the source code. If this is done, en dash is still available by typing ==&.

[75] If this behavior is undesirable, it can be easily suppressed; see “PATCH AREA em dash” twice in the source code. If this is done, em dash is still available by typing --&. (Note that Word, and perhaps other software, also may change two hyphens to em dash.) Even without making this patch, the change from two hyphens to em dash is disabled when typing the beginning of an HTML comment manually (<!--) and is automatically undone when typing the > at the end of the comment (-->).

[76] If the automatic conversion of three periods to ellipses is undesirable, you can undo the change by typing BACKSPACE, as in the next line. It is also possible to change the table so that three periods are not automatically changed to ellipsis; see “PATCH AREA ellipsis” twice in the source code. If this is done, ellipsis is still available by typing ...& (three periods plus &). Note that Word (and perhaps other software) also may change three periods to ellipsis; this may produce results that are different from what are reported here.

[77] That is, ellipsis followed by BACKSPACE produces three periods.

Description	Available characters	What to type
marks for commercial property	© ℗ ® ℠ ™	(c)& (p)& (r)& TM&
pseudo-fractions	⅍ ⅏ ℅ ℆	a/c& a/s& c/o& c/u&
communications	℻ ℡	FAX& TEL&
units of temperature	° ℃ ℉ K	0# °C& °F& K&
units of measurement	℔ µ ℥	lb& m[CTRL+/] oz&
other abbreviatory symbols	№ ℞	No& Rx&

7 Numerical, mathematical, and currency symbols

Of the many mathematical symbols available in Unicode, the following provide a basic set of the ones most commonly-used, at least for linguistic work. The following trigger keys are used:

Type of symbol	Trigger character (typed afterward)
symbol containing a solidus (slash)	[CTRL+ALT+/]
symbol containing a low line	[CTRL+SHIFT+|]
general mathematical symbols	#
currency symbols	$

Description	Available characters	What to type
fractions	↉ ½ ⅓ ⅔ ¼ ¾ ⅕ ⅖ ⅗ ⅘ ⅙ ⅚ ⅛ ⅜ ⅝ ⅞ $^{n(n)*}/_{n(n)*}$ (any combination)	*superscript-digit(s)* / *digit(s)* e.g. 1[CTRL+^]/2 for ½ or 2[CTRL+^]3/25 for $^{23}/_{25}$.[78]
inequalities	≡	=# or =[CTRL+SHIFT+|]
	≠	=[CTRL+ALT+/]
	≤ ≥	<# ># or <[CTRL+SHIFT+|] >[CTRL+SHIFT+|]
operators	+ − × ·	+ -# x# .#
	÷	/# or :[CTRL+ALT+|]
	±	+# or +[CTRL+SHIFT+|]
	∓	±[CTRL+SHIFT+#]
	ƒ	f#
quantifiers	∀ ∃ ∄	A# E# E#[CTRL+ALT+/]
set notation	∅	0[CTRL+ALT+/] (zero, not Oh)
	∪ ∩	U# U#[CTRL+SHIFT+#]
	∈ ∉	e# e#[CTRL+ALT+/]
per mille, per ten thousand	‰ ‱	%0& %00&
currency symbols	¢ € ₣ ƒ £ ¤ ¥	c$ E$ F$ f$ L$ O$ Y$

When the trigger key is an ordinary printable character which follows the base character, you can get all combinations of base character, modified character and trigger character by typing the trigger twice or three times. Thus:

[78] Only the first digit in the fraction needs to be superscripted; after that, the rest will be superscripted or subscripted automatically. Whenever a better-looking composite character is available it will be substituted for the individual pieces.

Combination	What to type
€	E$
E$	E$$
€$	E$$$

When the trigger key uses CTRL or ALT, typing it a second time has no effect.

8 Arrows and other graphic symbols

Keying of arrows and other graphic symbols depends heavily on the DIRECTIONAL MODIFIER KEYS, which consist of CTRL+ALT and optionally SHIFT plus the regular cursor control arrows. These determine what direction the graphic is oriented. The ones with SHIFT are rotated 45° counterclockwise from those that do not use SHIFT.

Direction of arrow		Directional modifier (typed afterward)
upwards	↑	[CTRL+ALT+UP]
upwards and left	↖	[CTRL+ALT+SHIFT+UP]
leftwards	←	[CTRL+ALT+LEFT]
downwards and left	↙	[CTRL+ALT+SHIFT+LEFT]
downwards	↓	[CTRL+ALT+DOWN]
downwards and right	↘	[CTRL+ALT+SHIFT+DOWN]
rightwards	→	[CTRL+ALT+RIGHT]
upwards and right	↗	[CTRL+ALT+SHIFT+RIGHT]

These key combinations are indicated in the chart below as *dirMod*.

You can also rotate these graphic symbols counterclockwise using the same trigger key that is used to turn phonetic characters: [CTRL+SHIFT+#]. The only difference is that, when used with graphic symbols, this key turns them only 45° or 90° (depending on the symbol) rather than 180°. This option is indicated in the chart below, when available, by the words "(plus optional 45°/90° rotation)".

Description	Available characters	What to type	
arrow, single shaft, head at one end	↖↑↗ ← → ↙↓↘	| ^ *dirMod*	(plus optional 45° rotation)
arrow, double shaft, head at one end	⇖⇑⇗ ⇐ ⇒ ⇙⇓⇘	| | ^ dirMod	(plus optional 45° rotation)
arrow, single shaft, heads at both ends	↔ ↕	^ | ^ *dirMod*	(plus optional 90° rotation) (typing either leftwards or rightwards produces horizontal; upwards or downwards produces vertical)
arrow, double shaft, heads at both ends	⇔⇕	^ | | ^ *dirMod*	(plus optional 90° rotation) (typing either leftwards or rightwards produces horizontal; upwards or downwards produces vertical)
crop marks	⌟ ⌞ ⌝ ⌜	# *dirMod*	(plus optional 90° rotation)