

Hangul	유니코드	유니코드
Hanunoo		
Hebrew	יוניקוד	יוניקוד
Hebrew (pointed)	יוֹנִיקוּדַיִם	יוֹנִיקוּדַיִם
Hebrew (Yiddish)	יוניקאד	יוניקאד
Hiragana (Japanese)	ゆにこおど	
Katakana (Japanese)	ユニコード	ユニコード
Kannada	ಯುನಿಕೋಡ್	ಯುನಿಕೋಡ್
Khmer	យូនីកូដ	យូនីកូដ
Lao		
Latin	Unicode	Unicode
Latin (IPA)	'juːnɪˌkoːd	'juːnɪˌkoːd
Latin (Am. Dict.)	Ūnĭcōdĕ	Ūnĭcōdĕ
Limbu		
Linear B		
Malayalam	യൂനികോഡ്	യൂനികോഡ്
Mongolian		
Myanmar		
Ogham		
Old Italic		
Oriya	ୟୁନିକୋଡ଼	ୟୁନିକୋଡ଼
Osmanya		
Runic (Anglo-Saxon)	*N+ILFM	*N+ILFM
Shavian		
Sinhala	යුනිකොද්	
Syriac	ܝܘܢܝܩܘܕܝܡ	ܝܘܢܝܩܘܕܝܡ
Tagbanwa		
Tagalog		
Tai Le		

Tamil	யூனிகோட்	யூனிகோட்
Telugu	యూనీకోడ్	యూనీకోడ్
Thaana		
Thai	ยูนิโคด	ยูนิโคด
Tibetan (Dzongkha)	ཡུ་ནི་ཀོ་དྲ།	ཡུ་ནི་ཀོ་དྲ།
Ugaritic		
Yi		

Notes:

There are different ways to transcribe the word "Unicode", depending on the language and script. In some cases there is only one language that customarily uses a given script; in others there are many languages. The goal here is at a minimum to collect at least one transcription for each script in a language customarily written in that script, with more languages if possible. If the transcription is the same for multiple languages in a script, then a single representative language is used.

Still missing are transcriptions for the items above in **RED** (in at least one language). I would appreciate any other transcriptions, or corrections for the ones listed here. Send to mark3@macchiato.com, using the directions below:

• Supplying Missing Items

- Most Latin-script languages will follow the spelling, and change the pronunciation. For any that would not, it would be good to have the alternate spelling.
- For non-Latin scripts the goal is to match the English pronunciation — *not* spelling. Above is the [IPA](#) (in phonemic transcription) that should be matched as closely as possible (without sounding affected in the target language)
- Text would be best in either the UTF-8 text, or the code points in hex HTML. E.g. either of the following:
 - "Юникод"
 - "Юникод"
 - Note: for [supplementary characters](#), there should be one hex number per code point, not two [surrogates](#):
 - 𐀀 *not* �&xDC00;
- If you have a good font, I'd also appreciate a GIF. It should be **96 x 24** bits, with the text centered, in black on white (plus grays if smoothed).

• Other Comments

- Because some browsers won't handle the text, both text and GIF image are supplied. If you can't read the text columns, see [Display Problems](#).
- The Chinese versions (inc. Bopomofo) are translations, not transcriptions, since "transcription in Chinese is pretty lame" [J. Becker].

- There are other "translations" of Unicode that may be in use, such as the Vietnamese "Thống Nhất Mã".
 - For sample pages in different languages on the Unicode site, see [What is Unicode?](#)
 - Americans are not generally used to IPA, and find a variety of different systems in their dictionaries. This one leaves the base letters as they are, and uses diacritics for pronunciation.
 - **Etymology of *Unicode***
 - Coined by J. Becker. Not related to previous usages, such as:
 - A telegraphic code in which one word or set of letters represents a sentence or phrase; a telegram or message in this. (late 19th century, OED)
 - According to my references, the prefix "uni" is directly from Latin while the word "code" is through French.
 - The original Indo-European apparently would have been *oino-kau-do ("one strike give"): *kau apparently being related to such English words as: hew, haggle, hoe, hag, hay, hack, caudad, caudal, caudate, caudex, coda, codex, codicil, coward, incus, and Kovač (personal name: "smith").
 - I will leave the exact derivations to the exegetes, but I like the association with "haggle" myself.
 - **Contributions**
 - This draws on contributions or comments from:
 - Dixon Au
 - Joe Becker
 - Maurice Bauhahn
 - Abel Cheung
 - Peter Constable
 - Michael Everson
 - Christopher John Fynn
 - Michael Kaplan
 - George Kiraz
 - Abdul Malik
 - Siva Nataraja
 - Roozbeh Pournader
 - Jonathan Rosenne
 - Jungshik Shin
-