

UMing TW (UMing TW (and UMinG TW MBE) from uming.ttc 0.2.20080216.2)

This Traditional Chinese Serif font covers Traditional Chinese, Simplified Chinese, Cantonese, and most European languages written with a Latin alphabet, also Vietnamese. I find the Russian text unreadable because the letters are spaced too far apart. The Greek letters lack accents which are required for normal usage of Greek.

For variations of the Latin and Cyrillic alphabets I attempt to show the whole alphabet, followed by Article 1 of the UDHR (it is traditional!). For other scripts and languages I merely display the digits (if not common 0..9) and Article 1.

Sources:

http://unicode.org/udhr/assemblies/first_article_all.html

<http://www.omniglot.com/>

<https://en.wikipedia.org>

For Chinese it is impossible to know exactly which codepoints you are likely to encounter. I have started by assuming that older fonts are probably adequate for common usage and so I have ignored coverage of CJK Extension B and later when trying to determine if a font is likely to be adequate.

I am primarily concerned with informal postings online, but I have tried to estimate which items might be found in more formal texts.

This font omits two codepoints (U+44D8, U+4A3C) from the nanshanjiang paragraphs in zhlipsum which I selected for testing language coverage, and one (U+9B2C) from the xiangyu paragraphs I selected.

CJK languages use an ellipsis codepoint to shown omissions or tailing-off in speech. The codepoint is ideally placed at mid-level (most old CJK fonts do that if they do not cover non-English Latin languages), but in Unicode 14 the mid-level ellipsis (a mathematical symbol) was recommended for this purpose. This font correctly puts both ellipses at mid-level, although that might look odd if an ellipsis appears in text using a Latin alphabet.

The UKai and UMinG TTCs were a proof of concept to support Minnan and Hakka, but the project apparently stalled. In the various versions (CN, HK, TW and TWMBE) only 2 glyphs (U+4EE4 and U+9AA0) differ in the Han glyphs, and 13 Bopomofo glyphs differ between TW and TWMBE.

This document was prepared using Xe_{La}TeX and fontspec. Missing glyphs will be replaced by an empty box. This font has only a regular weight.

1 Latin Alphabet (European variants)

1.1 Catalan

A a, B b, C c, Ç ç, D d, E e, É é, È è, F f, G g, H h, I i, J j, K k, L l, (EL and l1 or L·L l·l are a different pronunciation to LL and ll), M m, N n, O o, Ó ó, Ò ò, P p, Q q, R r, S s, T t, U u, Ü ü, V v, W w, X x, Y y, Z z

Tots els éssers humans neixen lliures i iguals en dignitat i en drets. Són dotats de raó i de consciència, i han de comportar-se fraternalment els uns amb els altres.

1.2 Czech

A a, Á á, B b, C c, Č č, D d, Ď ě, E e, É é, Ě ě, F f, G g, H h, Ch ch, I i, Í í, J j, K k, L l, M m, N n, Ň ň, O o, Ó ó. P p, Q q, R r, Ř ř, S s, Š š, T t, ě ě, U u, Ú ú. Ů ů, V v, W w, X x, Y y, Ý ý, Z z, Ž ž

Všichni lidé rodí se svobodní a sobě rovní co do důstojnosti a práv. Jsou nadáni rozumem a svědomím a mají spolu jednat v duchu bratrství.

1.3 Danish

A a, B b, C c, D d, E e, F f, G g, H h, I i, J j, K k, L l, M m, N n, O o, P p, Q q, R r, S s, T t, U u, V v, W w, X x, Y y, Z z, Æ æ, Ø ø, Å å

Alle mennesker er født frie og lige i værdighed og rettigheder. De er udstyret med fornuft og samvittighed, og de bør handle mod hverandre i en broderskabets ånd.

1.4 Dutch

As well as the standard 26 letters, the IJ digraph IJ ij may be encountered.

Alle mensen worden vrij en gelijk in waardigheid en rechten geboren. Zij zijn begiftigd met verstand en geweten, en behoren zich jegens elkander in een geest van broederschap te gedragen.

1.5 French

A a, À à, Â â, Æ æ, Ä ä, B b, C c, Ç ç, D d, E e, É é, È è, Ê ê, Ë ë, F f, G g, I i, Î î, Ï ï, J j, K k, L l, M m, N n, (Ñ ñ may be seen on words of Spanish origin), O o, Ô ô, Œ œ, Ö ö, P p, Q q, R r, S s, T t, U u, Ù ù, û û, Ü ü, V v, W w, X x, Y y, Ÿ ŷ, Z z

Tous les êtres humains naissent libres et égaux en dignité et en droits. Ils sont doués de raison et de conscience et doivent agir les uns envers les autres dans un esprit de fraternité.

1.6 German

A a, Ä ä, B b, C c, D d, E e, F f, G g, H h, I i, J j, K k, L l, M m, N n, O o, Ö ö, P p, Q q, R r, S s, T t, U u, Ü ü, V v, W w, X x, Y y, Z z and ß (but ss in Switzerland)

Alle Menschen sind frei und gleich an Würde und Rechten geboren. Sie sind mit Vernunft und Gewissen begabt und sollen einander im Geist der Brüderlichkeit begegnen.

1.7 Hungarian

A a, Á á, B b, C c, Cs cs, D d, Dz dz, Dzs dzs, E e, É é, F f, G g, Gy gy, H h, I i, Í í, J j, K k, L l, Ly ly, M m, N n, Ny ny, O o, Ó ó, Ő ő, Ű ű, P p, Q q, R r, S s, Sz sz, T t, Ty ty, U u, Ú ú, Ū ū, Ű ű, V v, W w, X x, Y y, Z z, Zs zs

Minden emberi lény szabadon születik és egyenlő méltósága és joga van. Az emberek, ésszel és lelkiismerettel bírván, egymással szemben testvéri szellemben kell hogy viseltessenek.

1.8 Icelandic

A a, Á á, B b, D d, Ð ð, (eth is not used for the first letter of any word) E e, É é, F f, G g, H h, I i, Í í, J j, K k, L l, M m, N n, O o, P p, R r, S s, T t, U u, Ú ú, V v, X x, Y y, Ý ý, Þ þ, Æ æ, Ö ö

Hver maður er borinn frjálst og jafn öðrum að virðingu og réttindum. Menn eru gæddir vitsmunum og samvitzku, og ber þeim að breyta bróðurlega hverjum við annan.

1.9 Italian

A a, À à, B b, C c, D d, E e, É é, È è, F f, G g, H h, I i, Í í, Ì ì, Ï ï, L l, M m, N n, O o, Ó ó, Ò ò, P p, Q q, R r, S s, T t, U u, Ú ú, Û ù, V v, Z z

Tutti gli esseri umani nascono liberi ed eguali in dignità e diritti. Essi sono dotati di ragione e di coscienza e devono agire gli uni verso gli altri in spirito di fratellanza.

1.10 Latvian

A a, Ā ā, B b, C c, Č č, D d, E e, Ē ē, F f, G g, Ģ ģ, H h, I i, Ī ī, J j, K k, Ķ ķ, L l, Ļ ļ, M m, N n, Ņ ņ, O o, P p, R r, S s, Š š, T t, U u, Ū ū, V v, Z z, Ž ž

Visi cilvēki piedzimst brīvi un vienlīdzīgi savā pašcienā un tiesībās. Viņi ir apveltīti ar saprātu un sirdsapziņu, un viņiem jāizturas citam pret citu brālības garā.

1.11 Lithuanian

A a, Ą ą, B b, C c, Č č, D d, E e, Ė ė, F f, G g, H h, I i, Į į, Y y, J j, K k, L l, M m, N n, O o, P p, R r, S s, Š š, T t, U u, Ū ū, V v, Z z, Ž ž

Visi žmonės gimsta laisvi ir lygūs savo orumu ir teisėmis. Jiems suteiktas protas ir sąžinė ir jie turi elgtis vienas kito atžvilgiu kaip broliai.

1.12 Maltese

A a, B b, Ċ ċ, D d, E e, F f, Ġ ġ, G g, Għ għ, H h, Ħ ħ, I i, IE ie, J j, K k, L l, M m, N n, O o, P p, Q q, R r, S s, T t, U u, V v, W w, X x, Z z

In addition, final vowels with grave accents are found in some words of Italian origin: à, è, ì, ò. ù.

Il-bnedmin kollha jitwielldu ħielsa u ugwali fid-dinjità u d-drittijiet. Huma mogħnija bir-raġuni u bil-kuxjenza u għandhom iġibu ruhhom ma' xulxin bi spirtu ta' aħwa.

1.13 Northern Sami

A a, Á á, B b, C c, Č č, D d, Đ đ, E e, F f, G g, H h, I i, J j, K k, L l, M m, N n, Ŋ ŋ, O o, P p, R r, S s, Š š, T t, Ț ț, U u, V v, Z z, Ž ž

Buot olbmot leat rieggádan friddjan ja olmmošárvvu ja olmmošvuoigatvuodaidi dáfus. Sii leat jierbmalaš olbmot geain lea oamedovdu ja sii gálggaše leat dego vieljačagat.

1.14 Polish

A a, Ą ą, B b, C c, Ć ć, D d, E e, Ę ę, F f, G g, H h, I i, J j, L l, Ł ł, M m, N n, O o, Ó ó, P p, R r, S s, Ś ś, T t, U u, W w, Y y, Z z, Ż ż, Ź ź

The letters Q q, V v, X x may be seen in foreign words.

Wszyscy ludzie rodzą się wolni i równi pod względem swej godności i swych praw. Są oni obdarzeni rozumem i sumieniem i powinni postępować wobec innych w duchu braterstwa.

1.15 Portuguese

A a, Á á, Â â, Ã ã, À à, B b, C c, Ç ç, D d, E e, É é, Ê ê, F f, G g, H h, I i, Í í, J j, K k, L l, M m, N n, O o, Ó ó, Ô ô, Õ õ, P p, Q q, R r, S s, T t, U u, Ú ú, V v, W w, X x, Y y, Z z

Todos os seres humanos nascem livres e iguais em dignidade e em direitos. Dotados de razão e de consciência, devem agir uns para com os outros em espírito de fraternidade.

1.16 Romanian

A a, Ă ă, Â â, B b, C c, D d, E e, F f, G g, H h, I i, Î î, J j, K k, L l, M m, N n, O o, P p, Q q, R r, S s, Ș ș, T t, Ț ț, U u, V v, W w, X x, Y y, Z z

Because Unicode was slow to add the below comma forms Ș ș, Ț ț the incorrect use of cedillas in Ș ș, Ț ț is still sometimes encountered.

Toate ființele umane se nasc libere și egale în demnitate și în drepturi. Ele sunt înzestrate cu rațiune și conștiință și trebuie să se comporte unele față de altele în spiritul fraternității.

1.17 Serbo-Croat (Latin)

A a, B b, C c, Č č, Ć ć, D d, Dž Dž dž (DŽ Dž dž), Đ đ, E e, F f, G g, H h, I i, J j, K k, L l, LJ Lj lj (LJ Lj lj), M m, N n, NJ Nj nj (NJ Nj nj), O o, P p, R r, S s, Š š, T t, U u, V v, Z z, Ž ž

Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i svijješću i treba da jedno prema drugome postupaju u duhu bratstva.

1.18 Spanish

A a, Á á, B b, C c, D d, E e, É é, F f, G g, H h, I i, J j, K k, L l, M m, N n, Ñ ñ, O o, Ó ó, P p, Q q, R r, S s, T t, U u, Ú ú, Ü ü, V v, W w, X x, Y y, Z z

Todos los seres humanos nacen libres e iguales en dignidad y derechos y, dotados como están de razón y conciencia, deben comportarse fraternalmente los unos con los otros.

1.19 Turkish

A a, Â â, B b, C c, Ç ç, D d, E e, F f, G g, Ğ ğ, H h, İ i, I ı, Î î, J j, K k, L l, M m, N n, O o, Ö ö, P p, R r, S s, Ş ş, T t, U u, û û, Ü ü, V v, Y y, Z z

Bütün insanlar hür, haysiyet ve haklar bakımından eşit doğarlar. Akıl ve vicdana sahiptirler ve birbirlerine karşı kardeşlik zihniyeti ile hareket etmelidirler.

1.20 Welsh

A a, Á á, Â â, À à, Ä ä, B b, C c, D d, Dd dd, E e, É é, Ê ê, È è, Ë ë, F f, Ff ff, G g, Ng ng, H h, i i, Í í, Î î, Ï ï, L l, Ll ll, M m, N n, O o, Ó ó, Ô ô, Ò ò, Ö ö, P p, Ph ph, R r, Rh rh, S s, T t, Th th, U u, Ú ú, Û û, Ü ü, W w, Ẃ ẃ, Ŵ ŵ, W̃ w̃, Y y, Ý ý, Ŷ ŷ, Ÿ ŷ, ŷ ŷ In addition, the letters J j, K k, Q q, V v, X x, Z z are used — J appears to be becoming assimilated into Welsh, the others are used in technical terms although sometimes transliterations are used. But K was part of the Welsh language until the bible was translated, when it was replaced by C because ‘the printers did not have enough Ks’.

Genir pawb yn rhydd ac yn gydradd â’ i gilydd mewn urddas a hawliau. Fe’ u cynysgaeddir â rheswm a chydwybod, a dylai pawb ymddwyn y naill at y llall mewn ysbryd cymodlon.

2 Cyrillic Alphabet

2.1 Russian

А а, Б б, В в, Г г, Д д, Е е, Ё ё, Ж ж, З з, И и, Й й, К к, Л л, М м, Н н, О о, П п, Р р, С с, Т т, У у, Ф ф, Х х, Ц ц, Ч ч, Ш ш, Щ щ, Ъ ъ, Ы ы, Ь ь, Э э, Ю ю, Я я

Все люди рождаются свободными и равными в своем достоинстве и правах. Они наделены разумом и совестью и должны поступать в отношении друг друга в духе братства.

3 CJK scripts

3.1 Cantonese

人人生出嚟就係自由嘅，喺尊嚴同權利上一律平等。佢哋具有理性同良心，而且應該用兄弟間嘅關係嚟互相對待。

3.2 Simplified Chinese

人人生而自由，在尊严和权利上一律平等。他们赋有理性和良心，并应以兄弟关系的精神互相对待。

3.3 Traditional Chinese

In some of my files I use 'TW' as a shorthand for Traditional Chinese used in Taiwan. This is the Mandarin version of Article 1:

人人生而自由，在尊嚴和權利上一律平等。他們賦有理性和良心，並應以兄弟關係的精神互相對待。

However, the common dialect in Taiwan, Taiwanese Hokkien (臺語) a.k.a. Minnan, is different and was what the UKai and Uming fonts were hoping to support. This is its Article 1, and this font fits one less ideogram on the first line for this text when using XeLaTeX.

人皆生而自由；在尊嚴及權利上均各平等。人各賦有理性良知，誠應和睦相處，情同手足。

4 Other variants of the Latin alphabet

4.1 Vietnamese

The interesting features of Vietnamese are the use of horns on ‘o’ and ‘u’ to provide additional vowels, and the use of diacriticals to show the tone. The letters themselves are pre-composed, but displaying some of the individual tone markings requires combining accents.

A a, Ă ă, Â â, B b, C c, D d, Đ đ, E e, Ê ê, G g, H h, I i, K k, L l, M m, N n, O o, Ô ô, Ơ ơ, P p, Q q, R r, S s, T t, U u, Ư ư, V v, X x, Y y

The additional markings for tone are as follows.

If two descriptions, they are northern / southern :

A a, Ă ă, Â â, E e, Ê ê, I i, O o, Ô ô, Ơ ơ, U u, Ư ư, Y y mid level
` À à, Ằ ằ, Ẳ ẳ, È è, Ẻ ẻ, Ì ì, Ò ò, Ồ ồ, Ờ ờ, Ù ù, Ừ ừ, Ỡ ỡ low falling
˘ Ằ ằ, Ẳ ẳ, Ẵ ẵ, Ẻ ẻ, Ẽ ẽ, Ỉ ỉ, Ỏ ỏ, Ỡ ỡ, Ỡ ỡ, Ủ ủ, Ữ ữ, Ỡ ỡ mid falling / dipping
~ Ằ ằ, Ẳ ẳ, Ẵ ẵ, Ẻ ẻ, Ẽ ẽ, Ỉ ỉ, Ỏ ỏ, Ỡ ỡ, Ỡ ỡ, Ủ ủ, Ữ ữ, Ỡ ỡ glottalized rising / dipping
˘ Ằ ằ, Ẳ ẳ, Ẵ ẵ, Ẻ ẻ, Ẽ ẽ, Ỉ ỉ, Ỏ ỏ, Ỡ ỡ, Ỡ ỡ, Ủ ủ, Ữ ữ, Ỡ ỡ high rising
. Ằ ằ, Ẳ ẳ, Ẵ ẵ, Ẻ ẻ, Ẽ ẽ, Ỉ ỉ, Ỏ ỏ, Ỡ ỡ, Ỡ ỡ, Ủ ủ, Ữ ữ, Ỡ ỡ glottalized falling / low rising

Tất cả mọi người sinh ra đều được tự do và bình đẳng về nhân phẩm và quyền lợi. Mọi con người đều được tạo hóa ban cho lý trí và lương tâm và cần phải đối xử với nhau trong tình anh em.

5 Quotation-marks, etc.

In real-world usage, particularly when looking at news sites, the presence or absence of the correct quote characters will make a difference. As a coder, I use '...' or "...". but in good text there are variations and this font supports:

left and right single quotes	angle brackets	<...>
'...'	double angle brackets	«...»
left and right double quotes	corner brackets	「...」
"..."	white corner brackets	『...』
single turned comma,		
with right single quote		'...'
double low 9 and left double quotes		„ ... “
„ ... “		
guillemets		«...»
angle quotes		<...>

In Spanish languages, the inverted exclamation-mark ¡ and inverted question-mark ¿ are necessary.

6 Symbols

6.1 Currency Symbols

Many countries have symbols for their currencies. In alphabetic order:

₵ Cedi (Ghana)	₼ Manat (Azerbaijan)
¢ Cent (many places)	₦ Naira (Nigeria)
₡ Colon (Costa Rica, El Salvador)	£ Pound Sterling
₹ Cruzeiro (Brazil)	₽ Rouble (Russia)
\$ Dollar (many places)	₸ Tenge (Kazakhstan)
₫ Dong (Vietnam)	₺ Turkish Lira
€ Euro	₩ Won (Korea)
₴ Hryvnia (Ukraine)	¥ Yen (Japan)

6.2 Other symbols in text

© Copyright	¶ Pilcrow
‡ Dagger	® Registered
‡ Double Dagger	§ Section
‰ Per mille	™ Trademark