

All the general text, the header and the footer are in Liberation Sans

This file attempts to provide a comparison of how the different fonts will look. Unlike Latin or Cyrillic alphabets, I anticipate that each font will occupy the same number of lines, with line breaks at the same places, for the common text. Therefore, instead of ordering by size I am ordering these alphabetically by the names known to fontconfig.

While I was looking ahead to revising or updating fonts for other CJK languages it became clear to me that a lot of what is on the web regarding punctuation in CJK languages either assumes that everything for Chinese is Simplified Chinese (CN), is misleading for a European reader (mention of guillemets when what is meant is double angle quotes), or is specific to the requirements of a particular publisher.

As I tried to sort out language coverage for Japanese fonts I realised that I would eventually need files to cover Cantonese (Hong Kong). I found very little I could paste. For preparing the HK dummy text I decided to try to create my own text in the spirit of 'English as She is Spoke [O novo guia da conversação em português e inglês]'. I decided to try to use similar text to also cover CN and TW [Mandarin using traditional Chinese, as in Taiwan] and to attempt to identify glyphs which

differed.

For that I started by looking at Wikipedia and then tried to use Google Translate to see what it said (it uses so-called AI, i.e. a Large Language Model). Much of that did not use the codepoints shown in Wikipedia, or else gave rubbish translations back into English. I then added some quotations and also used a few phrases from Omniglot. In the end I have got something which I hope will be useful to those who can read these languages, as well as to those like me who only want to have all the text rendered.

I have assumed that very little online text will require characters outside the Basic Multilingual Plane, so my text does not include anything from CJK Extension B or newer.

Unfortunately, Google Translate now uses Large Language Models. In some of the Japanese translations I could see that it had included digits instead of ideographs which means those specific translations are clearly not usable in Japanese. I cannot read any of the CJK languages, so it seems very likely that there are other errors.

The examples show my best efforts at correct punctuation, but they do not show all the possible variants. As

I revise the fonts I will comment on punctuation for their supported languages. I am making a best effort in this example to show mid-level ellipsis, but for fonts such as Noto and its derivatives that is not possible.

I have then followed that standard paragraph with a different paragraph of random text for each font. Compared to my lipsum CN files from early October the examples are now longer, to better match the longer common paragraph. Unlike with alphabetic languages, almost any random CJK text can be translated to what appears to be semi-reasonable text.

For the second paragraphs I have started with text left over from the translations, then where necessary pasted short chunks from online new sites and Bing translations.

For the CN, TW and HK second paragraphs I will try to use the translations of the same text where possible - some parts are just too long in some languages.

In the second paragraphs I have slightly reduced the punctuation, the original 'lipsum' is distorted Latin text.

Both of these fonts are medium weight.

## Noto Sans CJK HK 3.020

我係一個乜都唔識嘅新手。敬酒！用腐蝕性汽水溶解喉排水管入面嘅脂肪。哥哥講完再見就搭巴士離開。我隻氣墊船裝滿晒鱔。芥藍，白菜，.....對健康好。我個寶寶離開咗我，我隻騾仔鬍線，係扑克遊戲入面輸咗我嘅錢。《咁多麻煩》係唔係佢咁嘅？甜品餐牌蜜糖班戟、忌廉甜甜圈、車厘子批。噴古龍水嚟掩飾氣味。向總督磕頭。我希望我聽咗我媽咪同我講嘅嘢。搞掂

時間係一個錯覺午餐時間係雙重嘅。我寧願開心過任何一日都係咁嘅。歷史重演首先係悲劇其次係鬧劇。唯一可以係山頂搵到嘅禪就係你帶上去嗰度嘅禪。所有優秀嘅嘢都同難得一様咁難。人唔可以講嘢就一定要沉默。我過咗我嘅生活好耐我知道佢點樣運作。魔杖揀咗個巫師。跳繩係禁止嘅。癲嘅第一個跡象係你手掌上面有毛第二個就係搵緊佢咁。我嘅親愛嘅就好似一朵紅玫瑰嘅。潮語者最多可

## Chiron Hei HK 3.034

我係一個乜都唔識嘅新手。敬酒！用腐蝕性汽水溶解喉排水管入面嘅脂肪。哥哥講完再見就搭巴士離開。我隻氣墊船裝滿晒鱔。芥藍，白菜，.....對健康好。我個寶寶離開咗我，我隻騾仔鬍線，係扑克遊戲入面輸咗我嘅錢。《咁多麻煩》係唔係佢咁嘅？甜品餐牌蜜糖班戟、忌廉甜甜圈、車厘子批。噴古龍水嚟掩飾氣味。向總督磕頭。我希望我聽咗我媽咪同我講嘅嘢。搞掂

毛巾係星際搭車者可以擁有嘅最有用嘅嘢。小心豹。你行幾慢都唔緊要只要你唔停就得。冇人會踩兩次同一條河因為嗰條河唔係同一條河而佢亦唔係同一個人。人係被譴責為自由。如果你想知一個男人係點就要好好睇下佢點對待佢嘅下等人而唔係同佢嘅同等人。畀個太妃糖個小朋友食。喺織機度旋轉紗線。戴條頸巾保暖。西打酒、芝士、多士、吞拿魚三文治。喺後面泊車。困惑將係我嘅墓誌銘