

Source:	CheonHyeong Sim (沈天珩)
Title:	JMJ-source related glyph issues
Date:	2025-02-18
Action:	To be considered by Japan NB and WG2

Background

Japan did a huge horizontal extension (36,422 characters) for CJKUI from Unicode 16.0, all of the extended characters were MJ-source.

日本はUnicode16.0から日中韓統合表意文字に巨大な水平拡張（36,422文字）を行い、拡張した文字はすべてMJソースであった。

The huge workload inevitably brings some issues – the issues are not caused by Japan itself but by our WG2 convenor, Michel Suignard, who generates the CodeCharts; but still, I request Japan to state a clear position on some questions in the following texts, in that case, the possible confusion for the font designers will be reduced.

巨大なワークロードは必然的にいくつかの問題をもたらす——日本自身ではなく、WG2の召集人のミシェル・スイニャール氏（CodeChartsを生成する人）によって引き起こした。それでもなお、日本に次のテキストでいくつかの質問について明確な立場を示すよう求めている。その場合、フォントデザイナーにとっての混乱が減らされる。

Main Issues

In November, 2024, I accidentally discovered that the font of 322 (out of 367) JK-source characters in CJK-ExtC are changed from HeiseiMincho to IPAmjMincho, and all the J-source characters using MS Mincho in the compatibility block are also changed to IPAmjMincho. The remaining 45 JK-source characters in CJK-ExtC are not included in Moji Jōhō, so the font was retained. Immediately I did a check of all those hundreds of characters to confirm if any glyph changes appear, and found 1 major glyph change and 5 minor glyph changes, respectively:

2024年11月、日中韓統合表意文字拡張CのJKソースの322文字（367文字のうち）のフォントが平成明朝からIPAmj明朝に変更され、日中韓互換漢字でMS明朝を使用していたすべてのJソースの文字もIPAmj明朝に変更されていることを偶然発見した。日中韓統合表意文字拡張CのJKソースの残りの45文字は文字情報に含まれていないため、フォントは保持されている。すぐに、グリフの変更があったかどうかを確認するために、これらの数百の文字をすべてチェックし、1つの主要なグリフの変更と5つの小さなグリフの変更を見つけた。それぞれは：



Fig.1-6 Multiple Versions of U+2AD06/2AF7A/2B0C7/2B4CA/2B684/FA67 from zi.tools

As far as I know, Unibook calls multiple fonts in a given order, I wondered if Michel mistakenly put IPAmjMincho before HeiseiMincho (the original JK-source font) when generating CodeCharts, so I reported this issue to him. In his reply, he said:

知っている限り、Unibookは一定の順序で複数のフォントを呼び出す。ミシェルがCodeChartsを生成するときIPAmj明朝を平成明朝（JKソースの最初のフォント）の前に間違って置いたかどうか疑問に思ったので、この問題を彼に報告した。彼は答えでこう言った：

...

A major complication in this area is that Japan NB, unlike other NBs, prefer not to update historic sources contributed to the standard. This means that the J source definition in both 10646 and Unicode code chart production is an absurdly complicated set of almost 300 lines of cascading and precise range definition for the J source (compared to

few lines or at most a couple dozen lines for other sources). When JMJ extensions were requested, I saw that as an opportunity to reduce some of that, and at least not making it worse. I had in fact hoped to make it much simpler by replacing the complicated current mess by mostly using JMJ with the few occurrences where it does not have J source.

It seems that I went a bit too far, ...

In other words, this is intended by Michel, but not a request from Japan. Also I personally feel like that Japan may not want such font changes no matter whether there are glyph changes or not, but I am not entirely sure anyway.

つまり、これはミシエルの意図であるが、日本からのリクエストではない。また、グリフの変更があるかどうかにもかかわらず、日本はそんなフォントの変更を望まないかもしれないと個人的に感じているが、とにかく全く確信しているのではない。

FA65 貝 154.12	贈 J3-7C3D ≡ 8D08 贈 ~ 8D08 FE00
FA66 走 162.0	、 入 J4-7969 ≡ 8FB6 入 ~ 8FB6 FE00
FA67 走 162.8	逸 J3-7C59 ≡ 9038 逸 ~ 9038 FE01
FA68 佳 172.11	難 J3-7D63 ≡ 96E3 難 ~ 96E3 FE00
FA69 音 180.11	響 J3-7D76 ≡ 97FF 響 ~ 97FF FE00

The above 6 glyphs are specially reverted in the CodeCharts of Unicode 17.0 Alpha Review. However, this led to an inconsistency that the characters from the same source are shown with the different fonts:

上記の6つのグリフは、Unicode17.0のアルファレビューのCodeChartsでわざと戻されている。しかし、これにより、同じソースの文字が異なるフォントで表示されているという不一致になった：

← Fig.7 U+FA65 through U+FA69 from U17 Alpha Review

We could see that, all the glyphs are shown with IPAmjMincho except for U+FA67, which is shown with MS Mincho. The similar case also exists in CJK-ExtC.

U+FA67を除くすべてのグリフはIPAmj明朝で表示されているが、U+FA67はMS明朝で表示されている。同じようなケースは日中韓統合表意文字拡張Cでも存在している。

I was busy preparing the Postgraduate Entrance Examination that time, so I did not have too much time to thoroughly investigate the JMJ-related issues. In late December, after finishing the exam, I checked the JMJ-related issues and discovered more issues.

当時は大学院入試の準備で忙しく、JMJ関連の問題を詳しく調べる時間はあまりなかった。12月下旬、試験を終えた後、JMJ関連の問題をチェックしたところ、さらに多くの問題が見つかった。

6688	U+654C	654C 支 66.6	敌 G0-3550	敌 T4-2D44		JMJ-012973	敌
6689	U+654E	654E 支 66.7	教 GE-3B36	教 T3-3636	教 K0-4E67	JMJ-012975	教
6690	U+6553	6553 支 66.7	斂 GE-3B38	斂 HB2-D5A7	斂 T2-3569	K2-376E	斂

Fig.8 [WG2 N5221](#) from Japan NB



Fig.9 U+654E from U17 Alpha Review (same as in U16)

Intuitively, we could see the glyph difference of U+654E in WG2 N5221 (as shown in Fig.8) and in CodeCharts (as shown in Fig.9). After checking the font from CodeCharts, I found that it is HeiseiMincho – the primary J-source font for URO. It contains U+654E, even at that time U+654E did not have a J-source. As mentioned above, Unibook calls multiple fonts in a given order, so U+654E is preferentially shown with HeiseiMincho than IPAmjMincho.

直感的に、WG2N5221 (図8) とCodeCharts (図9) でのU+654Eのグリフの違いが見られる。CodeChartsでフォントを確認したところ、それは日中韓統合表意文字 (URO) の最初のJソースのフォントであるHeiseiMinchoであることがわかる。当時U+654EにはJソースがなかったにもかかわらず、そのフォントにはU+654Eが含まれている。前述のように、Unibookは一定の順序で複数のフォントを呼び出すため、U+654EはIPAmj明朝よりも平成明朝で優先的に表示されている。

This issue involves 18 characters, of which U+654E is the only one with glyph differences:

この問題には18文字が関係しており、そのうちU+654Eだけがグリフの違いのある文字である：

Codepoint	4EFC	50F4	51EC	5759	5CF5	6085
HeiseiMincho	任	僣	夙	罍	崕	悅
IPAmjMincho	任	僣	夙	罍	崕	悅

Codepoint	6120	654E	663B	6AE2	6DF2	6DF8
HeiseiMincho	愠	教	昂	檄	澆	清
IPAmjMincho	愠	教	昂	檄	澆	清
Codepoint	7682	8807	8D76	9592	9751	9AD9
HeiseiMincho	皂	螞	赶	閒	青	高
IPAmjMincho	皂	螞	赶	閒	青	高

For this issue, I would like to suggest Japan to consider to change the font of these 18 characters to IPAmjMincho since they are JMJ-source characters. If Japan NB agrees with my suggestion, then this will become some work for Michel.

この問題に関しては、それらの18文字はJMJソースの文字であるため、日本にこれらの文字のフォントをIPAmj明朝に変更することを検討するよう提案したいと思う。もし日本が私の提案に同意すれば、これはミシエルの仕事になるつもりである。

On Dealing with the Font Inconsistencies

From my own perspective as an (*amateur*) font designer, starting from the reasons of the consistency, I originally intended to suggest to replace all the J-source glyphs by IPAmjMincho (unless not included in the Moji-Jōhō collection, totally 46 characters, specifically U+9FC5 and 45 characters in CJK-ExtC). This seems to be the viewpoint of Michel as well.

(アマチュア)のフォントデザイナーとしての私の観点から、一貫性の理由から、もともとすべてのJソースのグリフをIPAmj明朝に置き換えることを提案したいと思っていた(文字情報に含まれていない文字を除く、具体的にはU+9FC5と日中韓統合表意文字拡張Cの45文字、合計46文字)。これはミシエルのアイデアでもあるようである。

However, after checking all the glyph differences of non-JMJ-source characters between HeiseiMincho and IPAmjMincho, I discovered that there is a huge quantity – this will be listed in the appendix of this document for reference (there may be some omissions, but even though, the quantity is such huge).

しかし、JMJソース以外の文字の平成明朝とIPAmj明朝のグリフの違いをすべて確認したところ、膨大な量があることがわかる。参考のためにこの文書の付録にリストされている(漏れがあるかもしれないが、それでも量はそのように膨大である)。

In this case, I would like to provide three strategies to Japan NB for consideration:

この場合、日本に検討するための3つの戦略を備え付けたいと思う：

The Best Strategy: / 上策 :

To change the font of all the non-JMJ-source characters mistakenly changed to IPAmjMincho by Michel except for the compatibility ones back to HeiseiMincho, no matter whether there is a (major/minor) glyph change or not. This is the safest and could maintain maximum stability.

グリフの（主要な/小さな）変更があるかどうかに関係なく、ミシェルによるIPAmj明朝に間違って変更されたすべてのJMJソース以外の文字（日中韓互換漢字を除く）のフォントを平成明朝に戻す。これは最も安全であり、最大限の安定性を維持することができる。

For the compatibility characters, which are originally shown with MS Mincho – the only appearance of MS Mincho in CodeCharts, personally I feel that we do not have the necessity to change back, even for U+FA67 mentioned above, which is the only character with glyph differences between MS Mincho and IPAmjMincho – but still, I am just suggesting, and the final decision will be up to Japan. My reason is that, 常用漢字表 (*List of Commonly Used Kanjis*) uses IPAMincho to show the Kanjis, and in that table, U+FA67, as a variant of U+9038, the stroke 乚 also attaches 一:

日中韓互換文字、すなわちもともとMS明朝で表示されていた文字（CodeChartsでMS明朝の唯一の登場）については、前述のようにMS明朝とIPAmj明朝でグリフが異なる唯一の文字であるU+FA67についても、元に戻す必要はないと個人的には感じているが、あくまで提案しているだけで、最終決定は日本次第である。その理由は、『常用漢字表』では漢字の表示にIPA明朝が使用されており、その表ではU+FA67がU+9038の異体字として、筆画「乚」は「一」に付いているからである：



Fig.10 [常用漢字表 P13](#)

The Medium Strategy: / 中策 :

To change the font of all the J-source characters to IPAmjMincho (unless not included in the Moji-Jōhō collection). The advantage is that it can achieve maximum consistency, which is better over the long run, but macroscopically, the disadvantage is that, there would be suddenly too many glyph changes, wherein some of them even change back to the JIS90 style! This will be a huge job to check and fix those glyphs – I do not know why Moji-Jōhō collection choose to use the JIS90 style as the default glyph (and JIS2004 style as the IVD) for some characters, I am just describing the facts, see the appendix for reference.

すべてのJソースの文字のフォントをIPAmj明朝に変更する（文字情報に含まれていない文字を除く）。メリットは最大限の一致性を実現できることで、長期的にはより良いが、マクロ的に見ると、デメリットは、グリフの変更が突然多すぎ、ひいては一部の文字がJIS90スタイルに戻ってしまう!それらのグリフをチェックして修正するのは大変な作業

になりそうである。文字情報がどうして一部の文字に対してJIS90スタイルをデフォルトのグリフとして（そしてJIS2004スタイルをIVDとして）使用することを選んだ訳はわからず、ここではただ事実を述べているだけである（付録を参照）。

The Worst Strategy: / 下策 :

To maintain the existing circumstance. The easiest for Michel since he needs to do nothing in this case; but the most inconsistent for Japan NB, and the most confused for font designers.

現状を維持する。この場合、ミシェルにとっては何もする必要がないので最も簡単であるが、日本にとっては最も不一致になり、フォントデザイナーにとっても最も混乱を招いている。

(End of the strategies)

I hope that Japan could state a clear position. That will help a lot.

日本が明確な立場を示すことを願っている。それは大いに役立つであろう。

(End of the main body of the document; appendix from next page)

附録

イ、平成明朝とIPAmj明朝のグリフの違い（JMJソース以外）

左はコードポイント、中は平成明朝、右はIPAmjMincho。合計は336文字（漏れがあるかもしれない）。

注1：「关」（U+5173）と「复」（U+590D）はただ幅が違っている。J4ソースのは部件としての使い方なので比較的狭く、JMJソースのは漢字としての使い方なので比較的広い。書き方は全く同じ。

注2：「冤」（U+51A4）、「叛」（U+53DB）、「屠」（U+5C60）、「庖」（U+5E96）、「愈」（U+6108）、「杓」（U+6753）など、数多くの字は、平成明朝のグリフはJIS2004スタイルであることに対し、IPAmjMinchoのデフォルトのグリフはJIS90スタイル（JIS2004スタイルのはIVDになっている）である。

03404	𠂇	𠂇	04FB9	𠂇	𠂇	053D5	𠂇	𠂇	05C60	屠	屠	0632D	捩	捩
03441	𠂇	𠂇	05093	𠂇	𠂇	053DB	叛	叛	05C7C	𠂇	𠂇	0633A	捩	捩
0356E	𠂇	𠂇	05099	備	備	053DF	𠂇	𠂇	05D34	𠂇	𠂇	0633B	捩	捩
03905	𠂇	𠂇	050C7	𠂇	𠂇	054A1	𠂇	𠂇	05D47	𠂇	𠂇	0633D	捩	捩
03971	𠂇	𠂇	050ED	𠂇	𠂇	054AC	𠂇	𠂇	05D61	𠂇	𠂇	06428	捩	捩
03AF4	𠂇	𠂇	05133	𠂇	𠂇	054CE	𠂇	𠂇	05D76	𠂇	𠂇	06429	捩	捩
03B78	𠂇	𠂇	05173	关	关	054EE	𠂇	𠂇	05DCD	𠂇	𠂇	0643F	捩	捩
03C2D	𠂇	𠂇	051A4	冤	冤	0550C	𠂇	𠂇	05E1F	𠂇	𠂇	06442	捩	捩
03CBC	𠂇	𠂇	051C2	𠂇	𠂇	0553E	𠂇	𠂇	05E96	庖	庖	0646F	捩	捩
03CDF	𠂇	𠂇	051D1	𠂇	𠂇	05561	𠂇	𠂇	05EA8	𠂇	𠂇	064D1	捩	捩
03E72	𠂇	𠂇	051DE	𠂇	𠂇	055F6	𠂇	𠂇	05EEB	𠂇	𠂇	0651D	捩	捩
03EE8	𠂇	𠂇	05233	𠂇	𠂇	0564F	𠂇	𠂇	05EFB	𠂇	𠂇	0658A	捩	捩
040DD	𠂇	𠂇	0523C	𠂇	𠂇	056C0	𠂇	𠂇	05F21	𠂇	𠂇	065A7	捩	捩
044D4	𠂇	𠂇	05296	𠂇	𠂇	057CF	𠂇	𠂇	05F63	𠂇	𠂇	065D4	捩	捩
04535	𠂇	𠂇	05298	𠂇	𠂇	0580D	𠂇	𠂇	05F98	𠂇	𠂇	065FC	捩	捩
04576	𠂇	𠂇	052DC	𠂇	𠂇	0590D	复	复	06098	𠂇	𠂇	0661E	捩	捩
0459B	𠂇	𠂇	05307	𠂇	𠂇	059B3	𠂇	𠂇	06108	愈	愈	066C4	捩	捩
046D5	𠂇	𠂇	0533B	医	医	059FF	𠂇	𠂇	061EF	𠂇	𠂇	06701	捩	捩
04749	𠂇	𠂇	053AF	𠂇	𠂇	05A17	𠂇	𠂇	06241	𠂇	𠂇	06753	杓	杓
04E40	𠂇	𠂇	053C9	𠂇	𠂇	05A29	𠂇	𠂇	0630D	捩	捩	06756	杖	杖

068B4	梃	梃
06969	榎	榎
06975	極	極
06994	榔	榔
06A39	樹	樹
06A3A	樺	樺
06A9D	檝	檝
06AD5	欖	欖
06B17	櫺	櫺
06B21	次	次
06B3A	欺	欺
06B53	歡	歡
06B59	歛	歛
06B9B	殞	殞
06BA9	燦	燦
06BDA	毳	毳
06C0E	氈	氈
06C26	氦	氦
06C4A	汙	汙
06D28	洩	洩
06D31	洎	洎
06D8D	洵	洵
06D8E	涎	涎
06DE1	淡	淡
06E4A	湊	湊
06E52	涓	涓
06ECF	滙	滙
06F06	漆	漆
06F10	藜	藜
06F5B	潛	潛

06F68	濼	濼
06FCA	濺	濺
0701E	瀟	瀟
07026	瀦	瀦
07037	瀼	瀼
0703A	灑	灑
07044	灑	灑
0704A	灑	灑
07065	轟	轟
07086	炆	炆
0708E	炎	炎
0717D	煽	煽
0718C	烟	烟
071A0	熠	熠
071A2	燧	燧
071B1	熱	熱
071C1	燁	燁
0723A	爺	爺
0724C	牌	牌
0728D	犍	犍
072E1	狡	狡
07305	狝	狝
07337	猷	猷
073D3	玆	玆
07467	琫	琫
074D3	瓓	瓓
074E0	瓠	瓠
074F6	瓶	瓶
07511	甌	甌
07526	甦	甦

07562	畢	畢
075BC	疼	疼
07618	痿	痿
076A3	瞞	瞞
076AF	𦉳	𦉳
076B4	皴	皴
077B1	瞞	瞞
07819	砬	砬
0791F	礮	礮
079F0	称	称
07A17	稗	稗
07A94	窆	窆
07AB3	竄	竄
07AC8	竈	竈
07B73	筵	筵
07B75	筵	筵
07BAD	箭	箭
07BAF	箎	箎
07BDD	箐	箐
07C63	籟	籟
07D5A	紉	紉
07D91	網	網
07D96	緹	緹
07DF6	縵	縵
07E2B	縫	縫
07FD5	翕	翕
07FDB	脩	脩
07FEB	翫	翫
07FF0	翰	翰
0809E	𦉳	𦉳

080F6	胶	胶
08120	脛	脛
08171	腓	腓
0817F	腿	腿
0819D	膝	膝
08292	芒	芒
082A0	苳	苳
0832D	苳	苳
08384	苳	苳
0839A	苳	苳
0839B	苳	苳
083AC	菟	菟
08415	芥	芥
08532	蔻	蔻
085AF	薯	薯
085C2	藜	藜
085F7	藜	藜
08654	虔	虔
086F8	蛸	蛸
08711	蛸	蛸
08713	蛸	蛸
08755	蝕	蝕
08805	蠅	蠅
08848	蚶	蚶
08869	𦉳	𦉳
0890E	𦉳	𦉳
08974	欄	欄
08975	榻	榻
089B4	覬	覬
08A1D	訝	訝

08A24	訖	訖
08A80	誦	誦
08A9F	諄	諄
08AB9	誹	誹
08AC7	談	談
08AEC	誓	誓
08AEE	諮	諮
08B2C	謬	謬
08B41	誨	誨
08B7E	譟	譟
08CC7	資	資
08D73	赳	赳
08DAF	趯	趯
08E4B	蹋	蹋
08E9D	躑	躑
08F2F	輯	輯
08F43	輳	輳
08FD4	返	返
0903C	逼	逼
09074	遴	遴
09080	邀	邀
090C3	郤	郤
090DE	郤	郤
09108	郟	郟
09139	郟	郟
0914B	酋	酋
09175	醇	醇
091AB	醫	醫
092CB	鋌	鋌
092CC	鋌	鋌

0939A	鎚	鎚
093F5	鑊	鑊
09453	鎚	鎚
09459	鑊	鑊
0946D	鑊	鑊
095A2	閔	閔
095BA	閔	閔
095D1	閔	閔
095D2	閔	閔
095D5	閔	閔
095DF	閔	閔
09622	阢	阢
096D2	阢	阢
09761	靡	靡
09771	鞞	鞞
09795	鞞	鞞
097AC	鞞	鞞
097AD	鞞	鞞

097B8	鞞	鞞
097BE	鞞	鞞
097E1	鞞	鞞
09832	頰	頰
098C7	颯	颯
098F4	飴	飴
09910	餐	餐
0995E	饒	饒
099C1	駁	駁
099EC	駟	駟
099F0	駟	駟
09A19	騙	騙
09A38	騙	騙
09A4A	驊	驊
09AB9	駁	駁
09ACE	膠	膠
09AD5	骸	骸
09B27	鬧	鬧

09B2D	鬪	鬪
09BAB	鮫	鮫
09BD6	鯖	鯖
09C76	鱈	鱈
09CF9	鴿	鴿
09D09	鴿	鴿
09D60	鴿	鴿
09DB2	鴿	鴿
09DDA	鴿	鴿
09E10	鶴	鶴
09E8E	農	農
09E9B	麇	麇
09E9D	麇	麇
09EDF	黠	黠
09EEE	黠	黠
09F69	齧	齧
09F6E	齧	齧
09F9C	龜	龜

09F9E	鼈	鼈
09FA1	鯁	鯁
09FA3	鯁	鯁
09FA5	鯁	鯁
0F929	朗	朗
0FA15	熙	熙
2000B	丈	丈
2231E	廻	廻
233E4	校	校
23465	梲	梲
25ED8	粼	粼
25FD4	絞	絞
269DD	涎	涎
26FF8	蝥	蝥
282F3	軛	軛
28A1E	銛	銛
29EC4	鰈	鰈
2AD06	櫛	櫛

2AF7A	瘵	瘵
2B0C7	糗	糗
2B4CA	鍛	鍛
2B684	鱻	鱻
2B764	奏	奏
2B7CF	菟	菟
2C11F	樺	樺
2C424	韭	韭
2C73C	莠	莠
2C82F	蟬	蟬
2C932	咎	咎
2CAF1	鎗	鎗
2CD77	鱻	鱻
2CD7A	鱻	鱻

ロ、文字情報に含まれていないJソースの文字

左はコードポイント、右は平成明朝。合計は46文字（漏れがないと確信）。

09FC5	禱
2A708	阡
2A70C	阡
2A784	𪗇
2A789	𪗇
2A8A6	𪗇
2A8B1	𪗇
2A915	𪗇

2A959	𪗇
2AA46	𪗇
2AA63	𪗇
2AA8D	𪗇
2AB23	𪗇
2ABAC	𪗇
2AC0C	𪗇
2ACAC	𪗇

2AD1C	𪗇
2AD25	𪗇
2AD61	𪗇
2AE09	𪗇
2AF3F	𪗇
2AF76	𪗇
2AFD3	𪗇
2AFFC	𪗇

2B016	𪗇
2B076	𪗇
2B07D	𪗇
2B0E5	𪗇
2B1C5	𪗇
2B26B	𪗇
2B2C1	𪗇
2B2E5	𪗇

2B2EC	𪗇
2B31D	𪗇
2B354	𪗇
2B3F2	𪗇
2B499	𪗇
2B4A0	𪗇
2B51B	𪗇
2B5DB	𪗇

2B674	𪗇
2B6B0	𪗇
2B6BA	𪗇
2B6C7	𪗇
2B6C9	𪗇
2B6CA	𪗇