Title: New ordering rules for Hangul, including rules for historic characters, and a DPRK delta

Source: Kent Karlsson
Date: 2001-10-30

Status: Expert Contribution
Document Type: Working Group Document
Action: For consideration by the UTC, JTC 1/SC 2/WG 2’s ad hoc on Korean, and JTC 1/SC 22/WG 20

Hangul syllables

A lot of Hangul syllables have a character of their own in the range (AC00-D7A3). They each have a canonical decomposition into two (choseong, jungseong) or three (choseong, jungseong, jongseong) Hangul Jamo characters. The Hangul Jamo characters are either basic or cluster characters. The Hangul syllables can thus be represented as sequences of Hangul Jamo characters.

The Hangul syllables alone can represent most modern Hangul texts. They cannot represent historic Hangul texts, or Hangul texts that use unusual consonant or vowel clusters. The latter have to use sequences of Hangul Jamo characters for their representation.
Basic and cluster Hangul Jamo characters
Cluster Hangul Jamo represent either clusters of two or three consonants, or clusters of two or three vowels. Cluster Jamo characters are for most (not all) occurring consonant and vowel clusters. One can also represent the sequence of consonants or sequence of vowels using basic Hangul Jamo characters, even though there is no decomposition of the cluster Jamos into basic Jamos (though there should be cross references). Unusual vowel or consonant clusters have to be represented using sequences of basic Hangul Jamo. The cluster Hangul Jamos are thus not needed.

A basic Hangul Jamo character represents one of the Hangul letters, or a historic variant of such a letter.

Composition of a Hangul syllable
A Hangul syllable has the following syntax:

```
Hsyllable::= C* CVsyl F* | C* CVFsyll F* | C+ V+ F*
```

where CVsyl is a precomposed consonants-vowels syllable, CVFsyll is a precomposed consonants-vowels-consonants syllable, C is an initial consonants Jamo, F is a final consonants Jamo, and V is a vowels Jamo.

However, there are at most (in total, however represented) three basic consonants in a consonant cluster, and at most (in total, however represented) three basic vowels in a vowel cluster.

Note the encoding trick here, to be able to determine syllable boundaries: the consonants are encoded twice, initial and final.

Question: How are the combining characters for Hangul tone marks,

```
302E HANGUL SINGLE DOT TONE MARK
302F HANGUL DOUBLE DOT TONE MARK
```

intended to be used? Do they apply to individual (basic) Jamos, or to syllables? How does either interpretations of those interact with normalisation?

Consider a Hangul syllable followed by a combining Hangul tone mark. Is that expected use? If so, if the string is turned into NFD form, the tone mark will apply only to (the last) jongseong. That seems strange! Or is the expected use of the tone marks with individual basic Jamos only? But if I V F T (where T is a tone mark) is turned into NFC, it will be S T (Hangul syllable, tone mark), so
the tone mark appears to apply to the syllable. If the tone mark only applies to Hangul jungseong (vowel) one still has the same problem for I V T --> S T. Or how are the Hangul tone marks intended to be used? My guess is that they apply to the syllable as a whole. But that's just my guess. If I'm right, then ideally, I think that the jongseong and jungseong should be combining characters, but I guess that may not be changeable by now.

End question.

**Alternative (better?) composition of a Hangul syllable**

A Hangul syllable could alternatively have the following, simpler, syntax:

\[
\text{Hsyllable::=} \text{CVsyl} \text{ V}^* \text{ F}^* \mid \text{CVFsyl} \text{ F}^* \mid \text{C} \text{ F}^* \text{ V}^+ \text{ F}^*
\]

The idea here is that the C characters (the choseong) are base characters, and V (jungseong) and F (jongseong, but not necessarily final, just non-first) are seen as combining characters. This would be more in line with how similar combinations are handled for other scripts in Unicode/10646. Further, even though the nominal weighting of cluster Jamos in the collation weighting table below would change as an implication of this (through generation of the table from <sort> decompositions in accordance with the above), there would be no change in the ordering of Hangul syllables from this change of aspect.

Whether the jungseong and jongseong characters are formally made combining or not is less important, and may be formally prohibited at this time. However, there seems to be no architectural hindrance to doing this, since (a) the formal decompositions of the cluster Jamos were removed from Unicode, and (b) the general combination from basic Jamos is not yet established in any of the Unicode standard and the ISO/IEC 10646 standard.

For this alternative approach to work (well), each basic Hangul choseong has to have a jongseong counterpart. For some variant letters, these counterparts are missing, namely:

| Hangul Jongseong Chitueumsios |
| Hangul Jongseong Ceongchieumsios |
| Hangul Jongseong Chitueumcieuc |
| Hangul Jongseong Ceongchieumcieuc |
| Hangul Jongseong Chitueumchieuch |
| Hangul Jongseong Ceongchieumchieuch |
| (Hangul Jongseong Filler) |
Hangul Jamo characters suggested by DPR of Korea

DPR of Korea appears to be working on a suggestion for including the following characters (given WG20/N867, and to some extent earlier documents):

- HANGUL CHOSEONG HYEOPATAKSORI NIEUN
- HANGUL CHOSEONG HYEOPATAKSORI TIKEUT (HYEOPATAKSORI TIEUT)
- HANGUL CHOSEONG MAREUMMIEUM
- HANGUL CHOSEONG HYEOPATAKSORI THIEUTH
- HANGUL CHOSEONG NIEUN-IEUNG
- HANGUL CHOSEONG HYEOPATAKSORI SSANGTIKEUT (HYEOPATAKSORI TU TIEUT)
- HANGUL CHOSEONG YESIEUNG-SIOS (EOKEUMISORI-SIEUS)

- HANGUL JUNGSEONG O-A
- HANGUL JUNGSEONG YO-A
- HANGUL JUNGSEONG YO-EO
- HANGUL JUNGSEONG U-YEO
- HANGUL JUNGSEONG U-EU
- HANGUL JUNGSEONG EU-A
- HANGUL JUNGSEONG YU-AE (YU-A-I)
- HANGUL JUNGSEONG YU-EO-I

Out of these, each of the following is a variant of a basic Jamo:

- HANGUL CHOSEONG HYEOPATAKSORI NIEUN (variant of NIEUN, long stem)
- HANGUL CHOSEONG HYEOPATAKSORI TIKEUT (variant of TIKEUT, long stem)
- HANGUL CHOSEONG MAREUMMIEUM (variant of MIEUM, 45 degrees rotated)
- HANGUL CHOSEONG HYEOPATAKSORI THIEUTH (variant of THIEUTH, long stem)

If sufficient supporting documentation about each these characters are provided, they should be encoded. If, in addition, the alternative combining approach described above is accepted, corresponding jongseong should then also be allocated:

- HANGUL JONGSEONG HYEOPATAKSORI NIEUN
- HANGUL JONGSEONG HYEOPATAKSORI TIKEUT
- HANGUL JONGSEONG MAREUMMIEUM
- HANGUL JONGSEONG HYEOPATAKSORI THIEUTH

Each of the remaining suggested characters is a cluster Jamo:

- HANGUL CHOSEONG NIEUN-IEUNG (cluster: <CHOSEONG NIEUN, CHOSEONG/JONGSEONG IEUNG>)
- HANGUL CHOSEONG HYEOPATAKSORI SSANGTIKEUT (cluster: <CHOSEONG HYEOPATAKSORI TIKEUT, CHOSEONG/JONGSEONG HYEOPATAKSORI TIKEUT>)
- HANGUL CHOSEONG YESIEUNG-SIOS (cluster: <CHOSEONG YESIEUNG, CHOSEONG/JONGSEONG SIOS>)
HANGUL JUNGSEONG O-A
(cluster: <JUNGSEONG O, JUNGSEONG A>; also: JUNGSEONG WA)
HANGUL JUNGSEONG YO-A (cluster: <JUNGSEONG YO, JUNGSEONG A>)
HANGUL JUNGSEONG YO-EO (cluster: <JUNGSEONG YO, JUNGSEONG EO>)
HANGUL JUNGSEONG U-YEO (cluster: <JUNGSEONG U, JUNGSEONG YEO>)
HANGUL JUNGSEONG U-EU (cluster: <JUNGSEONG U, JUNGSEONG EU>)
HANGUL JUNGSEONG YU-AE (YU-A-I)
(cluster: <JUNGSEONG YU, JUNGSEONG A, JUNGSEONG I>)
HANGUL JUNGSEONG YU-EO-I
(cluster: <JUNGSEONG YU, JUNGSEONG EO, JUNGSEONG I>)

These need not be encoded, since they can be represented by sequences of non-cluster Jamos. However HANGUL JUNGSEONG O-A appears to already be allocated, by the name HANGUL JUNGSEONG WA (though the proposal may actually turn out to be for a variant letter).

**Suggested allocation to code positions (if the characters are well documented and then accepted)**

115A HANGUL CHOSEONG HYEOPATAKSORI NIEUN
115B HANGUL CHOSEONG HYEOPATAKSORI TIKEUT
115C HANGUL CHOSEONG MAREUMMIEUM
115D HANGUL CHOSEONG HYEOPATAKSORI THIEUTH
115E (unused)

11A3 HANGUL JONGSEONG HYEOPATAKSORI NIEUN (combining counterpart)
11A4 HANGUL JONGSEONG HYEOPATAKSORI TIKEUT (combining counterpart)
11A5 HANGUL JONGSEONG MAREUMMIEUM (combining counterpart)
11A6 HANGUL JONGSEONG HYEOPATAKSORI THIEUTH (combining counterpart)
11A7 (HANGUL JONGSEONG FILLER) (combining counterpart)

11FA HANGUL JONGSEONG CHITUEUMSIOS (combining counterpart)
11FB HANGUL JONGSEONG CEONGCHIEUMSIOS (combining counterpart)
11FC HANGUL JONGSEONG CHITUEUMCIEUC (combining counterpart)
11FD HANGUL JONGSEONG CEONGCHIEUMCIEUC (combining counterpart)
11FF HANGUL JONGSEONG CEONGCHIEUMCHIEUCH (combining counterpart)

**Suggestions for the Hangul part of the ISO/IEC 14651 CTT**

In order to keep 14651 and UTS 10 in synchrony, corresponding changes to UTS 10 are also suggested.
Current (ISO/IEC 14651:2001) ordering for Hangul handles ‘modern’ Hangul well, provided that the text is represented in such a way that a syllable is just a single precomposed Hangul syllable, or composed of a ‘modern’ Initial, ‘modern’ Vowel, and optionally ‘modern’ Final Hangul Jamo, where the Jamos may be cluster Jamos. It does not handle ‘historic’ Hangul Jamo characters well, nor does it handle well Hangul syllables where the consonant or vowel clusters are composed from multiple single Hangul Jamo letters according to the general syntax above. This proposal is intended to remedy this, by handling Hangul in a way that is similar to how other alphabetic scripts are handled. The result is that ‘historic’ Hangul characters (and compositions for those) are ordered among the “modern” Hangul letters as expected, and that compositions from basic Jamo letters are ordered as expected. The ordering here has not yet been thoroughly reviewed by experts in Hangul, so there may be changes.

%%% BEGIN SUGGESTED HANGUL EXCERPT FOR 14651 CTT

% Declaration of collating symbols for Hangul (similar to how this is done for other alphabetic scripts) (order here is arbitrary):

% Vowels: 11 basic, plus 'filler'
collating_symbol <S1160> % HANGUL JUNGSEONG FILLER
collating_symbol <S1161> % HANGUL JUNGSEONG A
collating_symbol <S1163> % HANGUL JUNGSEONG YA
collating_symbol <S1165> % HANGUL JUNGSEONG EO
collating_symbol <S1167> % HANGUL JUNGSEONG YEO
collating_symbol <S1169> % HANGUL JUNGSEONG O
collating_symbol <S116D> % HANGUL JUNGSEONG YO
collating_symbol <S116E> % HANGUL JUNGSEONG U
collating_symbol <S1172> % HANGUL JUNGSEONG YU
collating_symbol <S1173> % HANGUL JUNGSEONG EU
collating_symbol <S1175> % HANGUL JUNGSEONG I
collating_symbol <S119E> % HANGUL JUNGSEONG ARAEA (ARAI)

% Initial consonants: 17 basic, plus 'filler'
collating_symbol <S115F> % HANGUL CHOSEONG FILLER
collating_symbol <S1100> % HANGUL CHOSEONG KIYEOK (KIEUK)
collating_symbol <S1102> % HANGUL CHOSEONG NIEUN
collating_symbol <S1103> % HANGUL CHOSEONG TIKEUT (TIEUT)
collating_symbol <S1105> % HANGUL CHOSEONG RIEUL
collating_symbol <S1106> % HANGUL CHOSEONG MIEUM
collating_symbol <S1107> % HANGUL CHOSEONG PIEUP
collating_symbol <S1109> % HANGUL CHOSEONG SIOS (SIEUS)
collating_symbol <S1140> % HANGUL CHOSEONG PANSIOS (PANISORI)
collating_symbol <S110B> % HANGUL CHOSEONG IEUNG
collating_symbol <S114C> % HANGUL CHOSEONG YESIEUNG (EOKEUMISORI)
collating_symbol <S110C> % HANGUL CHOSEONG CIEUC (JIEUJ)
collating_symbol <S110E> % HANGUL CHOSEONG CHIEUCH
% Final (or non-first) consonants: 17 basic

% Weighting of collating symbols for Hangul (the order here is important):

% Vowels: 11 basic, plus 'filler'

% Initial consonants: 17 basic, plus 'filler'
% In addition, `<SFFFF>` (declared elsewhere) is used to make vowels ordered *after* the final (or non-first) consonants, % when they are *not* the first vowel in the cluster. Further, the `<VRNTn>` collation symbols are used for letter variants.

% Weighting table for Hangul, except Hangul syllables (which are collated by their NFD form) and also except
% Hangul compatibility characters (which will be included when this table is generated by the “sifter” program).
% The order here is arbitrary (except for the fourth level weight, which is unimportant),
% but the order used here is, for review purposes, the one implied by the weights.

% Vowels (lightest):
% If, and only if, the current match to the collating element is preceded, in NFD form,
% (immediately! or with Hangul tone mark between?!) by a Hangul jungseong (vowel),
% the weights with a '?' after them are included. This is in order to make non-first
% vowel characters in a vowel sequence heavier than final (or non-first) consonants,
% which in turn is needed to get the expected ordering for Hangul syllables.

<U1160> "<SFFFF>?<S1160>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1160> % HANGUL JUNGSEONG FILLER

<U1161> "<SFFFF>?<S1161>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1161> % HANGUL JUNGSEONG A
.<U1165> "<SFFFF>?<S1165>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1165> % HANGUL JUNGSEONG EO
.<U1169> "<SFFFF>?<S1169>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1169> % HANGUL JUNGSEONG O
.<U1176> "<SFFFF>?<S1176>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1176> % HANGUL JUNGSEONG A-O
.<U1180> "<SFFFF>?<S1180>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1180> % HANGUL JUNGSEONG O-EQ
.<U1181> "<SFFFF>?<S1181>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1181> % HANGUL JUNGSEONG O-E
.<U1182> "<SFFFF>?<S1182>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1182> % HANGUL JUNGSEONG O-YE
.<U1184> "<SFFFF>?<S1184>"; "<BASE>?<BASE>"; "<MIN>?<MIN>"; <U1184> % HANGUL JUNGSEONG O-YA

9
% HANGUL JUNGSEONG YO-YAE (YO-YA-I)
% HANGUL JUNGSEONG YO-YEO
% HANGUL JUNGSEONG YO-O
% HANGUL JUNGSEONG YO-I
% HANGUL JUNGSEONG U-AE (U-A-I)
% HANGUL JUNGSEONG WEO (U-EO)
% HANGUL JUNGSEONG U-EO-EU
% HANGUL JUNGSEONG WE (U-EO-I)
% HANGUL JUNGSEONG U-U
% HANGUL JUNGSEONG WI (U-I)
% HANGUL JUNGSEONG YU-A
% HANGUL JUNGSEONG YU-EO
% HANGUL JUNGSEONG YU-E (YU-EO-I)
% HANGUL JUNGSEONG YU-U
% HANGUL JUNGSEONG YU-I
% HANGUL JUNGSEONG EU-U
% HANGUL JUNGSEONG EU-EU
% HANGUL JUNGSEONG EU-I
% HANGUL JUNGSEONG SSANGARAEA
% HANGUL JUNGSEONG ARAEA (ARAIA)
% HANGUL JUNGSEONG ARAEA-EO
% HANGUL JUNGSEONG ARAEA-U
% HANGUL JUNGSEONG ARAEA-I
% HANGUL JUNGSEONG ARAEA-AE (ARAIA-AE)
% HANGUL JUNGSEONG SSANGARAEA

10
<U115F> <S115F>; <BASE>; <MIN>; <U115F> % HANGUL CHOSEONG FILLER

<U1100> <S1100>; <BASE>; <MIN>; <U1100> % HANGUL CHOSEONG KYEOK (KIEUK)
<U1101> "<S1100><S1100>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1101> % HANGUL CHOSEONG SSANGKYEOK

<U1102> <S1102>; <BASE>; <MIN>; <U1102> % HANGUL CHOSEONG NIEUN
% HANGUL CHOSEONG HYEOPATAKSORI NIEUN
<U1113> "<S1102><S1100>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1113> % HANGUL CHOSEONG NIEUN-KYEOK

<U1114> "<S1102><S1102>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1114> % HANGUL CHOSEONG SSANGNIEUN

<U1115> "<S1102><S1103>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1115> % HANGUL CHOSEONG NIEUN-TIKEUT

<U1116> "<S1102><S1107>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1116> % HANGUL CHOSEONG NIEUN-PIEUP

<U1103> <S1103>; <BASE>; <MIN>; <U1103> % HANGUL CHOSEONG TIEUT (TIEUK)
% HANGUL CHOSEONG HYEOPATAKSORI TIEUT
<U1117> "<S1103><S1100>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1117> % HANGUL CHOSEONG TIEUT-KYEOK

<U1104> "<S1103><S1103>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1104> % HANGUL CHOSEONG SSANGTIEUT

<U1105> <S1105>; <BASE>; <MIN>; <U1105> % HANGUL CHOSEONG RIEUL
<U1118> "<S1105><S1102>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1118> % HANGUL CHOSEONG RIEUL-NIEUN

<U1119> "<S1105><S1105>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1119> % HANGUL CHOSEONG SSANGRIEUL

<U111B> "<S1105><S110B>"; "<BASE><BASE>"; "<MIN><MIN>"; <U111B> % HANGUL CHOSEONG KAPYEOUNRIEUL

<U111A> "<S1105><S1112>"; "<BASE><BASE>"; "<MIN><MIN>"; <U111A> % HANGUL CHOSEONG RIEUL-HIEUH

<U1106> <S1106>; <BASE>; <MIN>; <U1106> % HANGUL CHOSEONG MIEUM
% HANGUL CHOSEONG MAREUMMIEUM
<U111C> "<S1106><S1107>"; "<BASE><BASE>"; "<MIN><MIN>"; <U111C> % HANGUL CHOSEONG MIEUM-PIEUP

<U111D> "<S1106><S1110>"; "<BASE><BASE>"; "<MIN><MIN>"; <U111D> % HANGUL CHOSEONG KAPYEOUNMIEUM

<U1107> <S1107>; <BASE>; <MIN>; <U1107> % HANGUL CHOSEONG PIEUP
<U111E> "<S1107><S1100>"; "<BASE><BASE>"; "<MIN><MIN>"; <U111E> % HANGUL CHOSEONG PIEUP-KYEOK

<U111F> "<S1107><S1102>"; "<BASE><BASE>"; "<MIN><MIN>"; <U111F> % HANGUL CHOSEONG PIEUP-NIEUN

<U1120> "<S1107><S1103>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1120> % HANGUL CHOSEONG PIEUP-TIKEUT

<U1108> "<S1107><S1107>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1108> % HANGUL CHOSEONG SSANGPIEUP

<U112C> "<S1107><S1110>"; "<BASE><BASE>"; "<MIN><MIN>"; <U112C> % HANGUL CHOSEONG KAPYEOUNSSANGPIEUP

<U1121> "<S1107><S1110>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1121> % HANGUL CHOSEONG PIEUP-SIOS

<U1122> "<S1107><S1109>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1122> % HANGUL CHOSEONG PIEUP-SIOS-KYEOK

<U1123> "<S1107><S1109><S1103>"; "<BASE><BASE><BASE>"; "<MIN><MIN><MIN>"; <U1123> % HANGUL CHOSEONG PIEUP-SIOS-TIKEUT

<U1124> "<S1107><S1109><S1107>"; "<BASE><BASE><BASE>"; "<MIN><MIN><MIN>"; <U1124> % HANGUL CHOSEONG PIEUP-SIOS-PIEUP

<U1125> "<S1107><S1109><S1109>"; "<BASE><BASE><BASE>"; "<MIN><MIN><MIN>"; <U1125> % HANGUL CHOSEONG PIEUP-SSANGSIOS

<U1126> "<S1107><S1109><S110C>"; "<BASE><BASE><BASE>"; "<MIN><MIN><MIN>"; <U1126> % HANGUL CHOSEONG PIEUP-SIOS-CIEUC

<U1128> "<S1107><S110B>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1128> % HANGUL CHOSEONG KAPYEOUNPIEUP

<U1127> "<S1107><S110C>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1127> % HANGUL CHOSEONG PIEUP-CIEUC

<U1129> "<S1107><S1110>"; "<BASE><BASE>"; "<MIN><MIN>"; <U1129> % HANGUL CHOSEONG PIEUP-CHIEUCH

<U112A> "<S1107><S1111>"; "<BASE><BASE>"; "<MIN><MIN>"; <U112A> % HANGUL CHOSEONG PIEUP-THIEUTH

<U1109> <S1109>; <BASE>; <MIN>; <U1109> % HANGUL CHOSEONG SIOS (SIEUS)
% HANGUL CHOSEONG KIYEOK (KIEUK)
% HANGUL CHOSEONG JONGSEONG KIYEOK (RIEUK)
% HANGUL CHOSEONG JONGSEONG NIEUN (TIEUT)
% HANGUL CHOSEONG JONGSEONG RIEUL (PIEUP)
% HANGUL CHOSEONG JONGSEONG KAPYEOUNPHIEUP
% HANGUL CHOSEONG JONGSEONG KIEUK
% HANGUL CHOSEONG JONGSEONG PHIEUP-PIEUP
% HANGUL CHOSEONG JONGSEONG KAPYEOUNPHIEUP
% HANGUL CHOSEONG JONGSEONG YEORINGHIEUH (MOKKUMEONGTHEOCHIMSORI)
% Final (or non-first) consonants:
<U11B3> "<S11AF><S11BA>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11B3> % HANGUL JONGSEONG RIEUL-SIOS
<U11D6> "<S11AF><S11BA><S11BA>"; "<BASE><BASE><BASE>"; "<MIN><MIN><MIN>"; <U11D6> % HANGUL JONGSEONG RIEUL-SSANGSIOS
<U11D7> "<S11AF><S11EB>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11D7> % HANGUL JONGSEONG RIEUL-PANSIOS
<U11DB> "<S11AF><S11BF>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11DB> % HANGUL JONGSEONG RIEUL-KHIEUKH
<U11E1> "<S11AF><S11C2>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11E1> % HANGUL JONGSEONG RIEUL-HIEUH
<U11F0> "<S11AF><S11C2>"; "<BASE><VRNT1>"; "<MIN><MIN>"; <U11F0> % HANGUL JONGSEONG RIEUL-YEORINHIEUH
</U11B7> <S11B7>; <BASE>; <MIN>; <U11B7> % HANGUL JONGSEONG MIEUM
<U11DA> "<S11B7><S11A8>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11DA> % HANGUL JONGSEONG MIEUM-KIYEOK
<U11DE> "<S11B7><S11AF>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11DE> % HANGUL JONGSEONG MIEUM-RIEUL
<U11EE> "<S11B7><S11B8>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11EE> % HANGUL JONGSEONG MIEUM-PIEUP
<U11BB> "<S11B8>; <BASE>; <MIN>; <U11BB> % HANGUL JONGSEONG PIEUP
<U11F1> "<S11B8><S11BA>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11F1> % HANGUL JONGSEONG PIEUP-SIOS
<U11E6> "<S11B8><S11BF>"; "<BASE><BASE>"; "<MIN><MIN>"; <U11E6> % HANGUL JONGSEONG PIEUP-KHIEUKH
</U11BD> <S11BD>; <BASE>; <MIN>; <U11BD> % HANGUL JONGSEONG CIEUC (JIEUJ)
<U11BE> "<S11BE>; <BASE>; <MIN>; <U11BE> % HANGUL JONGSEONG CHIEUCH (JHIEUJ)
<U11BF> "<S11BF>; <BASE>; <MIN>; <U11BF> % HANGUL JONGSEONG KHEUKH
Hangul compatibility letters (non-conjoining), circled, parenthesised, and halfwidth Hangul letters are not covered above but will be when (if...) the table is generated by the “sifter” program. (For the tailoring below, those characters, for the cases needing tailoring, are included.)

Suggested tailoring for DPR Korea ordering
The suggestion below only covers the known differences for “modern” Hangul, and only extends these known differences to variants and near variants of “modern” Hangul letters. It does not take into account collation differences for “historic” Hangul as given in earlier DPRK collation delta suggestions, as they appear to be neither well established, nor systematic.

% Begin suggested DPR Korea delta for collation of Hangul.

% Vowels (these vowel clusters are ordered after I)
reorder_after <S1175> % HANGUL JUNGSEONG I

% Declaration of extra collating symbols:
collating_symbol <S1162> % HANGUL JUNGSEONG AE (A-I)
collating_symbol <S1164> % HANGUL JUNGSEONG YAE (YA-I)
collating_symbol <S1166> % HANGUL JUNGSEONG E (EO-I)
collating_symbol <S1168> % HANGUL JUNGSEONG YE (YEO-I)
% Weighting of them (after the weighting of I as per the reorder statement above):
<S1162>  % HANGUL JUNGSEONG AE (A-I)
<S1164>  % HANGUL JUNGSEONG YAE (YA-I)
<S1166>  % HANGUL JUNGSEONG E (EO-I)
<S1168>  % HANGUL JUNGSEONG YE (YEO-I)
<S116C>  % HANGUL JUNGSEONG OE (O-I)
<S1171>  % HANGUL JUNGSEONG WI (U-I)
<S1174>  % HANGUL JUNGSEONG YI (EU-I)

<S116A>  % HANGUL JUNGSEONG WA (O-A)
<S116F>  % HANGUL JUNGSEONG WEO (U-EO)
<S116B>  % HANGUL JUNGSEONG WAE (O-A-I, O-AE)
<S1170>  % HANGUL JUNGSEONG WE (U-EO-I, U-E)

<S116A>  % HANGUL JUNGSEONG WA (O-A)
<S116F>  % HANGUL JUNGSEONG WEO (U-EO)
<S116B>  % HANGUL JUNGSEONG WAE (O-A-I, O-AE)
<S1170>  % HANGUL JUNGSEONG WE (U-EO-I, U-E)

% Initial consonants (these consonant clusters, and IEUNG/YESIEUNG, are ordered after YEORINHIEUH):
reorder_after <S1159>  % HANGUL CHOSEONG YEORINHIEUH

% Declaration of extra collating symbols:
collating_symbol <S1110>  % HANGUL CHOSEONG SSANGKIYEOK (TOINKIEUK)
collating_symbol <S1104>  % HANGUL CHOSEONG SSANGTIKEUT (TOINTIEUT)
collating_symbol <S1108>  % HANGUL CHOSEONG SSANGPIEUP (TOINPIEUP)
collating_symbol <S110A>  % HANGUL CHOSEONG SSANGSIOS (TOINSIEUS)
collating_symbol <S110D>  % HANGUL CHOSEONG SSANGCIEUC (TOINIJIEUJ)

<S1101>  % HANGUL CHOSEONG SSANGKIYEOK (TOINKIEUK)
<S1104>  % HANGUL CHOSEONG SSANGTIKEUT (TOINTIEUT)
<S1108>  % HANGUL CHOSEONG SSANGPIEUP (TOINPIEUP)
<S110A>  % HANGUL CHOSEONG SSANGSIOS (TOINSIEUS)
<S110D>  % HANGUL CHOSEONG SSANGCIEUC (TOINIJIEUJ)
<S110B>  % HANGUL CHOSEONG IEUNG
<S114C>  % HANGUL CHOSEONG YESIEUNG (EOKEUMISORI)

% Final (or non-first) consonants (likewise)
reorder_after <S11F9>  % HANGUL JONGSEONG YEORINHIEUH
collating_symbol <S11A9>  % HANGUL JONGSEONG SSANGKIYEOK (TOINKIEUK)
collating_symbol <S11AE_S11AE>  % HANGUL JONGSEONG SSANGTIKEUT (TOINTIEUT)
collating_symbol <S11B8_S11B8>  % HANGUL JONGSEONG SSANGPIEUP (TOINPIEUP)
collating_symbol <S11BB_S11BB>  % HANGUL JONGSEONG SSANGSIOS (TOINSIEUS)
collating_symbol <S11BD_S11BD>  % HANGUL JONGSEONG SSANGCIEUC (TOINIJIEUJ)
% Declare collating elements for sequences that are to be collated as units:

collating_element <U1161_U1175> from "<U1161><U1175>" % HANGUL JUNGSEONG AE (A-I), composed
collating_element <U1163_U1175> from "<U1163><U1175>" % HANGUL JUNGSEONG YAE (YA-I), composed
collating_element <U1165_U1175> from "<U1165><U1175>" % HANGUL JUNGSEONG E (EO-I), composed
collating_element <U1167_U1175> from "<U1167><U1175>" % HANGUL JUNGSEONG YE (YEO-I), composed
collating_element <U1169_U1175> from "<U1169><U1175>" % HANGUL JUNGSEONG OE (O-I), composed
collating_element <U116E_U1175> from "<U116E><U1175>" % HANGUL JUNGSEONG WI (U-I), composed
collating_element <U1173_U1175> from "<U1173><U1175>" % HANGUL JUNGSEONG YI (EU-I), composed

collating_element <U1169_U1161> from "<U1169><U1161>" % HANGUL JUNGSEONG WA (O-A), composed
collating_element <U116E_U1165> from "<U116E><U1165>" % HANGUL JUNGSEONG WEO (U-EO), composed

collating_element <U1169_U1161_U1175> from "<U1169><U1161><U1175>" % HANGUL JUNGSEONG WAE (O-A-I), composed1
collating_element <U1169_U1162> from "<U1169><U1162>" % HANGUL JUNGSEONG WAE (O-A-I), composed2
collating_element <U116E_U1165_U1175> from "<U116E><U1165><U1175>" % HANGUL JUNGSEONG WEO (U-EO-I), composed1
collating_element <U116E_U1166> from "<U116E><U1166>" % HANGUL JUNGSEONG WEO (U-EO-I), composed2

collating_element <U1100_U1100> from "<U1100><U1100>" % HANGUL CHOSEONG SSANGKIYEOK, composed
collating_element <U1103_U1103> from "<U1103><U1103>" % HANGUL CHOSEONG SSANGTIKEUT, composed
collating_element <U1107_U1107> from "<U1107><U1107>" % HANGUL CHOSEONG SSANGPIEUP, composed
collating_element <U1109_U1109> from "<U1109><U1109>" % HANGUL CHOSEONG SSANGSIOS, composed
collating_element <U113C_U113C> from "<U113C><U113C>" % HANGUL CHOSEONG CHITUEUMSSANGSIOS, composed
collating_element <U113E_U113E> from "<U113E><U113E>" % HANGUL CHOSEONG CEONGCHIEUMSSANGSIOS, composed
collating_element <U113C_U113E> from "<U113C><U113E>" % HANGUL CHOSEONG CHITUEUMSSANGSIOS, composed

collating_element <U110C_U110C> from "<U110C><U110C>" % HANGUL CHOSEONG SSANGCIJEUC, composed
collating_element <U114E_U114E> from "<U114E><U114E>" % HANGUL CHOSEONG CHITUEUMSSANGCIJEUC, composed
collating_element <U1150_U1150> from "<U1150><U1150>" % HANGUL CHOSEONG CEONGCHIEUMSSANGCIJEUC, composed

reorder after <U1100> % HANGUL CHOSEONG KIYEOK
% (exact position does not matter, but at least put them among the other Hangul)
As before, the weights that are marked with a '?' are included if and only if the preceding (looked upon in NFD form) character was a Hangul jungseong (vowel).

% Initial consonants:

% Vowels:

% HANGUL JUNGSEONG AE (A-I)

% HANGUL JUNGSEONG YAE (YA-I)

% HANGUL JUNGSEONG E (EO-I)

% HANGUL JUNGSEONG YE (YEO-I)

% HANGUL JUNGSEONG OE (O-I)

% HANGUL JUNGSEONG WI (U-I)

% HANGUL JUNGSEONG YI (EU-I)

% HANGUL JUNGSEONG WAE (O-A-I)

% HANGUL JUNGSEONG WEO (U-EO-I)

% HANGUL JUNGSEONG PIEUP-SSANGSIOS

% HANGUL JUNGSEONG SSANGKIYEOK

% HANGUL JUNGSEONG SSANGTIKEUT
% HALFWIDTH HANGUL LETTER SSANGTIKEUT
% HANGUL CHOSEONG HYEOPATAKSORI SSANGTIKEUT, composed

% HANGUL CHOSEONG SSANGPIEUP
% HANGUL CHOSEONG SSANGPIEUP, composed
% HANGUL CHOSEONG SSANGSIOS
% HANGUL CHOSEONG SSANGSIOS, composed
% HANGUL CHOSEONG SSANGCIEUC
% HANGUL CHOSEONG SSANGCIEUC, composed

% Final (or non-first) consonants:
% HANGUL JONGSEONG RIEUL-SSANGSIOS
% HANGUL JONGSEONG MIEUM-SSANGSIOS
% HANGUL JONGSEONG SSANGKIYEOK
% HANGUL JONGSEONG SSANGKIYEOK, composed

reorder_end    % end of proposed delta for DPR Korea ordering of Hangul.
References

**ISO/IEC 10646-1:2000**

**Unicode 3.0**
The Unicode standard, version 3.0. (In particular the compatibility decompositions in the UCD and (algorithmic) canonical decomposition of Hangul syllables.)

**UCD 3.1**
Unicode character database, version 3.1.

**ISO/IEC 14651:2001**
International string ordering and comparison – Method for comparing character strings and description of the common template tailorable ordering.

**UTS 10**
Unicode technical standard 10, Unicode collation algorithm.

**ISO/IEC JTC1 N5999**
DPRK, Proposal for a New Work Item, Amendment to the part concerning Korean characters in ISO/IEC 10646-1:1993. (NWI rejected.)

**ISO/IEC JTC1/SC2/WG2 N2243**
DPRK, Proposal for the addition of 14 Korean Alphabets to 10646-1.

**ISO/IEC JTC1/SC2/WG2 N2246**
DPRK, Evidence for arrangement of Korean characters.

**ISO/IEC JTC1/SC2/WG2 N2249R**
Kent Karlsson, Draft 14651 CTT tailoring delta for D.P.R. Korea.

**ISO/IEC JTC1/SC22/WG20 N858**
(Same as SC2/WG2 N2249R.)

**ISO/IEC JTC1/SC22/WG20 N867**
DPRK, Proposed ISO/IEC 14651 CTT tailoring delta for D P R of Korea.

**ISO/IEC JTC1/SC22/WG20 N876**
KIM, Kyongsok, Comments regarding N867, the DPRK proposed delta file for IS 14651

**ISO/IEC JTC1/SC22/WG20 N8nn**
Kent Karlsson, Revised draft ISO/IEC 14651:2001 CTT tailoring delta for D.P.R. Korea. Review of N867. (Not posted?)