Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation

Title: <u>Ordering rules for Hangul</u>

Source: Kent Karlsson
Date: 2001-11-29
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

1   Introduction

The Hangul script as such is very elegantly designed. However, its incarnation in 10646/Unicode is far from elegant. This paper is about restoring the elegance of Hangul, as much as it can be restored, for the process of string ordering.

1.1   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 in the ranges 1100-1112, 1161-1175, and 11A8-11C2. The choseong are leading consonants, one of which is mute. The jungseong are vowels. And the jongseong are trailing consonants. A Hangul Jamo character is either a letter or letter cluster.

The Hangul syllable characters alone can represent most modern Hangul words. They cannot represent historic Hangul words (Middle Korean), nor modern/future Hangul words using syllables not preallocated. However, all Hangul words can elegantly be represented by sequences of single-letter Hangul Jamo characters plus optional tone mark.

1.2 Single-letter and cluster Hangul Jamo characters

Cluster Hangul Jamo characters represent either clusters of two or three consonants, or clusters of two or three vowels. Cluster Jamo characters for most (not all) *occurring* consonant and vowel clusters are allocated. One can also represent the sequence of consonants or sequence of vowels using single-letter Hangul Jamo characters, even though there is no longer any *formal* decomposition of the cluster Jamos into single-letter Jamos. All vowel or consonant clusters can be represented using sequences of single-letter Hangul Jamos. The cluster Hangul Jamos are thus not needed, especially not the historic ones that do not figure in the decomposition of Hangul syllable characters. A single-letter Hangul Jamo character represents a basic Hangul letter, or a variant of such a letter. A possible problem is that the variant consonant letters are only allocated as choseong, with no jongseong counterpart. There may also be a small number of missing variant letters. However, no more cluster Jamos should be allocated. Note also that some vowel combinations look very much the same (like e.g. A-I and I-EO, EU-YO and YU-EU), but for each such same-looking pair of vowels apparently only one is allowed in Korean.

1.3 Composition of Hangul syllables

A Hangul syllable has the following syntax (see page 53 of The Unicode Standard version 3.0, with adjustment for tone marks and for what normalisation to normal form C (NFC) may produce):

Hangul-syllable ::=   L+ V+ T* M* | L* LVsyllablechar V* T* M* | L* LVTsyllablechar T* M*

where LVsyllablechar is a precomposed consonants-vowels syllable character, LVTsyllablechar is a precomposed consonants-vowels-consonants syllable character, L is a leading consonants Jamo or an Lf (a choseong filler), T is a trailing consonants Jamo, V is a vowels Jamo or a Vf (a jungseong filler), and M is any combining mark, in particular a Hangul tone mark [U+302E, U+302F]. (Is any Cf character allowed inside a Hangul syllable sequence?) The tone mark (if any, or more generally, the combining sequence) applies to the entire preceding syllable, not just the last part of it, since the Hangul syllable components, including the precomposed Hangul syllable characters, are conjoining characters, not base characters. The tone mark glyphically appears at the left of a syllable, so for a L V T M syllable, where M is a Hangul tone mark, the glyph for M is to be rendered to the left of the (possibly composed) glyph for L V T, not to the left of the (sub)glyph for T. Hangul syllables in NFD form are simpler:

Hangul-syllable-NFD ::=   L+ V+ T* M*.

Ideally, the use of cluster Jamo characters should be completely avoided. Note that Hangul is a (rather elegantly designed) alphabetic script, and typographic features, such as cluster ligatures, variant (sub)glyph selection, and syllable layout should be handled by font mechanisms. Ordering of Hangul syllables should be based on a weighting scheme that orders cluster characters as sequences of letters. At one point the cluster Jamos had compatibility decompositions into single-letter Jamos. Ideally, as is done below, when possible the cluster Hangul Jamos should be treated as if they had canonical decompositions into the corresponding sequence of single-letter Hangul Jamos.

In practice, there are at most (in total, however represented) three single-letter consonants (but transliterating words like "sturskt" or "västkustskt" may overrun that...) in a consonants cluster, and at most (in total, however represented) three single-letter vowels in a vowel cluster. Virtual, i.e. no change to the backing store, filler characters may be introduced automatically, in case the given Hangul string otherwise would not be parseable into Hangul syllables (see page 53 of The Unicode Standard version 3.0). <u>The two Hangul filler characters should normally not be explicitly used.</u>

Note the encoding trick used for the Hangul Jamos, to be able to determine syllable boundaries: the consonants are encoded twice, leading and trailing. Other ways that could have been used would include (a) using a separator character, or (b) using combining characters for the Hangul letters following the first one in a syllable (somewhat similar to the approach chosen); compare the combining Latin letters above that have recently been encoded.

Actually, something similar to the separator character approach is to some extent represented in Unicode. The Hangul compatibility letters and half-width letters, which encode the consonants only once each, have a different kind of FILLER character, 3164 HANGUL FILLER (and FFA0 HALFWIDHT HANGUL FILLER), which when converting from a sequence of Hangul compatibility letters (or half-width compatibility letters) to proper conjoining Hangul Jamo letters, should (in most occasions, see however the grammar below) be interpreted as switching from trailing to leading consonants (the compatibility mapping to Hangul Jungseong filler is irrelevant); vowels (and HANGUL FILLER, see grammar below) also mark the switch from leading to trailing consonants. Note that the choseongness or jongseongness of the compatibility mappings are irrelevant. Thus, when converting from Hangul compatibility letter sequences to proper conjoining Hangul letters, the Hangul compatibility syllables have the syntax:

Hangul-compatibility-syllable ::=   H (C+|H) (W+|H) (C+|H) M*

where C is a (possibly half-width) Hangul consonant(s) letter, W is a (possibly half-width) Hangul vowel(s) letter, and H is a (possibly half-width) FILLER character (a minimal amount of FILLERs can partially be automatically inserted to fullfill that grammar). It may be reasonable to convert, say, keyboard input in compatibility Hangul to a string of conjoining Jamos (H (C+|H) (W+|H) (C+|H) M* → L+ V+ T* M*). Compatibility Hangul is not expected to display as conjoined Hangul syllables, but display as a sequence of free-standing letters.

Another possibility is to consider each (half-width) Hangul compatibility letter as an isolated form (C → Lf Vf T or L Vf; and W → Lf V), but that leaves HANGUL FILLER useless and has an indeterminate mapping of the consonants. Further, if a Hangul compatibility character is really isolated, and converted according to the first scheme (and while any cluster characters are expanded!), the insertion of (virtual) choseong or jungseong fillers will then be done as usual with conjoining Jamo. That way we are getting in principle the same end result for *isolated* Hangul compatibility characters for both approaches.

Since the first approach is more general, it should be preferred. The first approach is therefore what is used for Hangul compatibility characters in the ordering rules below. This is easy enough for collation, but may be needlessly complicated elsewhere. Circled and parenthesised Hangul compatibility characters are ordered as freestanding.

2   Suggestions for the Hangul part of ISO/IEC 14651

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' leading, 'modern' vowel, and optionally 'modern' trailing 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 single-letter Jamos are ordered as expected. The ordering here has not yet been thoroughly reviewed by experts in Hangul, so there may be changes.

A difficulty is that the ordering of Hangul is cluster based. This is handled by having prehandling rules, that introduce a cluster end mark character (actually a non-character) at the end of each cluster. The cluster end mark character is given a weighting in such a way that the desired ordering is achieved: lighter than anything else. It may be that this prehandling is considered inappropriate for 14651 itself, but may instead be made a required part of tailorings that are intended to handle Hangul. Likewise for the weighted non-characters that are introduced by the prehandling; their weight declaration and weighting may be more suitable in tailorings intended to handle Hangul.

In order to keep ISO/IEC 14651 and UTS 10 in synchrony, corresponding changes to UTS 10 are also suggested.

2.1   Prehandling for cluster ordered Hangul

In order to get the correct ordering for all compositions of Hangul, the following prehandling steps must be done. These steps are done also for (half-width) compatibility Hangul letter sequences (see section 1.3 above).

1.   Canonical (and algorithmic) decomposition of Hangul syllable characters according to Unicode 2.0 (and later).

2.   (Optional) Replace each cluster Jamo with the corresponding sequence of single-letter Jamos.

3.   (Optional) Insertion of a minimal amount of choseong fillers and jungseong fillers or compatibility fillers to make the Hangul Jamo string (or Hangul compatibility letter string) parseable as Hangul syllables (according to L+ V+ T* M* or to H (C+|H) (W+|H) (C+|H) M*), see page 53 of The Unicode Standard 3.0.

4.   Insertion of a special cluster end character (used just for clustered collation) **B**, in the following way: L+ V+ T* M* → L+ **B** V+ **B** T* **B** M* and H (C+|H) (W+|H) (C+|H) M* → H (C+|H) **B** (W+|H) **B** (C+|H) **B** M*. (If step three is not done, this insertion must still be done at appropriate places, *as if* fillers had been inserted in step 3.)

The cluster end character **B** (U+FDD1 is used below) is weighted lighter than any letter of any script. The last step is needed to get the clustered ordering expected for Hangul. The overall ordering is thus as follows:

- FILLER characters are always ignored at the first three collation weighting levels.

- First Hangul vowel (via the special cluster end character, U+FDD1) in a vowel cluster.

- First Hangul consonant (via the special cluster end character, U+FDD1) in a Hangul syllable, except when at the very beginning of a Hangul string.

- First (non-ignored at the first level) character after a Hangul string (via the special cluster end character, U+FDD1).

- Non-first consonant (merging choseong and jongseong), or first consonant when at beginning of Hangul string.

- Vowels (jungseong).

B
L,C   T   V,W

As a small improvement, a U+FDD1 is not needed after the vowels, *if* the (non-filler) vowels are ordered after the trailing consonants: L+ V+ T* M* → L+ **B** V+ T* **B** M* and H (C+|H) (W+|H) (C+|H) M* → H (C+|H) **B** (W+|H) (C+|H) **B** M* (reason: there is no ordering requirement between the Ts and the **B**Ts, so they can collate identically, similarly for the Cs and the **B**Cs). See section 4 below for more improvements. This improvement is assumed below for the weighting of parenthesised and circled Hangul.

B
L,C   T
V,W
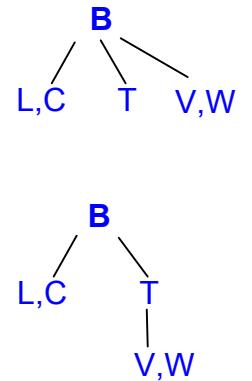
2.2 Suggested new Hangul excerpt for the 14651 CTT

%%% 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):

collating_symbol <B>        % FDD1, special cluster end character (for any script). Ordered before any letters or digits of any script.
```

```
% The 17 basic Hangul consonants:
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
collating_symbol <S110F>    % HANGUL CHOSEONG KHIEUKH
collating_symbol <S1110>    % HANGUL CHOSEONG THIEUTH
collating_symbol <S1111>    % HANGUL CHOSEONG PHIEUPH
collating_symbol <S1112>    % HANGUL CHOSEONG HIEUH
collating_symbol <S1159>    % HANGUL CHOSEONG YEORINHIEUH (MOKKUMEONGTHEOCHIMSORI)


collating_symbol <C2>       % FDD2, special cluster start character that may be used for two-consonant Hangul clusters.
collating_symbol <C3>       % FDD3, special cluster start character that may be used for three-consonant Hangul clusters.
collating_symbol <CI>       % FDD6, special cluster start character that may be used for leading ieung.


% The 11 basic Hangul vowels:
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 (ARAIA)


collating_symbol <V2>       % FDD4, special cluster start character that may be used for two-vowel Hangul clusters.
collating_symbol <V3>       % FDD5, special cluster start character that may be used for three-vowel Hangul clusters.


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


<B>             % FDD1, cluster end mark introduced by the required prehandling.
```

```
% The 17 basic Hangul consonants:
<S1100>        % ㄱ, HANGUL CHOSEONG KIYEOK (KIEUK) (k, g)

<S1102>        % ㄴ, HANGUL CHOSEONG NIEUN (n), and variant of nieun

<S1103>        % ㄷ, HANGUL CHOSEONG TIKEUT (TIEUT) (t, d), and variant of tikeut

<S1105>        % ㄹ, HANGUL CHOSEONG RIEUL (r, l)

<S1106>        % ㅁ, HANGUL CHOSEONG MIEUM (m), and variant of mieum

<S1107>        % ㅂ, HANGUL CHOSEONG PIEUP (p)

<S1109>        % ㅅ, HANGUL CHOSEONG SIOS (SIEUS) (s), and variants of sios

<S1140>        % ㅿ, HANGUL CHOSEONG PANSIOS (PANISORI)

<S110B>        % ㅇ, HANGUL CHOSEONG IEUNG (silent when leading, ng when trailing)

<S114C>        % ㆁ, HANGUL CHOSEONG YESIEUNG (EOKEUMISORI)

<S110C>        % ㅈ, HANGUL CHOSEONG CIEUC (JIEUJ) (j), and variants of cieuc

<S110E>        % ㅊ, HANGUL CHOSEONG CHIEUCH (ch)

<S110F>        % ㅋ, HANGUL CHOSEONG KHIEUKH (kh)

<S1110>        % ㅌ, HANGUL CHOSEONG THIEUTH (th), and variant of thieuth

<S1111>        % ㅍ, HANGUL CHOSEONG PHIEUPH (ph, f?)

<S1112>        % ㅎ, HANGUL CHOSEONG HIEUH (h)

<S1159>        % ㆆ, HANGUL CHOSEONG YEORINHIEUH (MOKKUMEONGTHEOCHIMSORI)
<C2>           % FDD2, non-character that may be introduced by the prehandling before two-letter Hangul consonants clusters
<C3>           % FDD3, non-character that may be introduced by the prehandling before three-or-more-letter Hangul consonants clusters
<CI>           % FDD6, non-character that may be introduced by the prehandling for leading Hangul ieung

% The 11 basic Hangul vowels:
<S1161>        % ㅏ, HANGUL JUNGSEONG A

<S1163>        % ㅑ, HANGUL JUNGSEONG YA

<S1165>        % ㅓ, HANGUL JUNGSEONG EO

<S1167>        % ㅕ, HANGUL JUNGSEONG YEO

<S1169>        % ㅗ, HANGUL JUNGSEONG O

<S116D>        % ㅛ, HANGUL JUNGSEONG YO

<S116E>        % ㅜ, HANGUL JUNGSEONG U

<S1172>        % ㅠ, HANGUL JUNGSEONG YU

<S1173>        % ㅡ, HANGUL JUNGSEONG EU

<S1175>        % ㅣ, HANGUL JUNGSEONG I

<S119E>        % ㆍ, HANGUL JUNGSEONG ARAEA (ARAIA)
<V2>           % FDD4, non-character that may be introduced by the prehandling before two-letter Hangul vowel clusters
<V3>           % FDD5, non-character that may be introduced by the prehandling before three-or-more-letter Hangul vowel clusters
```

% Further, the <VRNT1> and <VRNT2> second level collation symbols (weighted elsewhere in the CTT) are used for letter variants,
% and <COMPAT> and <NARROW> third level collation symbols (weighted elsewhere in the CTT) are used for compatibility variants.

% Weighting table for Hangul, except Hangul syllables (which are collated by their NFD form).
% 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 as assigned above.

% Hangul FILLERs are ignored (like control and format control characters, as well as punctuation):

<U115F> IGNORE;IGNORE;IGNORE;<U115F> % HANGUL CHOSEONG FILLER
<U1160> IGNORE;IGNORE;IGNORE;<U1160> % HANGUL JUNGSEONG FILLER
<U3164> IGNORE;IGNORE;IGNORE;<U3164> % HANGUL FILLER
<UFFA0> INGORE;IGNORE;IGNORE;<UFFA0> % HALFWIDTH HANGUL FILLER

% Hangul tone marks (handled as accents):

<U302E> IGNORE;<D302E>;<MIN>;<U302E> % HANGUL SINGLE DOT TONE MARK (applies to entire syllable)
<U302F> IGNORE;<D302F>;<MIN>;<U302F> % HANGUL DOUBLE DOT TONE MARK (applies to entire syllable)

% The cluster end mark has the lightest first level weight. Used for clustered collations:
<UFDD1> <B>;<BASE>;<MIN>;<UFDD1>   % FDD1, non-character that is introduced by the prehandling

% Hangul precomposed syllables are decomposed algorithmically in the required prehandling.

%<AC00>..<D7A3> "......<B>......<B>......<B>";"<BASE>.....<BASE>";"<MIN>.....<MIN>";........

% The 17 Hangul consonants, with clusters, compatibility forms, and variants (some variants not (yet) allocated):

<U1100> <S1100>;<BASE>;<MIN>;<U1100> % HANGUL CHOSEONG KIYEOK (KIEUK)
<U11A8> <S1100>;<BASE>;<MIN>;<U11A8> % HANGUL JONGSEONG KIYEOK (KIEUK)
<U3131> <S1100>;<BASE>;<COMPAT>;<U3131> % HANGUL LETTER KIYEOK
<UFFA1> <S1100>;<BASE>;<NARROW>;<UFFA1> % HALFWIDTH HANGUL LETTER KIYEOK
<U3200> "<S1100><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3200> % PARENTHESIZED HANGUL KIYEOK ("<S1100><B><B><B>")
<U3260> "<S1100><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3260> % CIRCLED HANGUL KIYEOK ("<S1100><B><B><B>")
<U320E> "<S1100><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U320E> % PARENTHESIZED HANGUL KIYEOK A
          ("<S1100><B><S1161><B><B>")

```
<U326E> "<S1100><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U326E> % CIRCLED HANGUL KIYEOK A
          ("<S1100><B><S1161><B><B>")
<U1101> "<S1100><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U1101> % HANGUL CHOSEONG SSANGKIYEOK
<U11A9> "<S1100><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U11A9> % HANGUL JONGSEONG SSANGKIYEOK
<U3132> "<S1100><S1100>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3132> % HANGUL LETTER SSANGKIYEOK
<UFFA2> "<S1100><S1100>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFA2> % HALFWIDTH HANGUL LETTER SSANGKIYEOK
<U11C3> "<S1100><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U11C3> % HANGUL JONGSEONG KIYEOK-RIEUL
<U11AA> "<S1100><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U11AA> % HANGUL JONGSEONG KIYEOK-SIOS
<U3133> "<S1100><S1109>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3133> % HANGUL LETTER KIYEOK-SIOS
<UFFA3> "<S1100><S1109>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFA3> % HALFWIDTH HANGUL LETTER KIYEOK-SIOS
<U11C4> "<S1100><S1109><S1100>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11C4> % HANGUL JONGSEONG KIYEOK-SIOS-KIYEOK

<U1102> <S1102>;<BASE>;<MIN>;<U1102> % HANGUL CHOSEONG NIEUN
<U11AB> <S1102>;<BASE>;<MIN>;<U11AB> % HANGUL JONGSEONG NIEUN
<U3134> <S1102>;<BASE>;<COMPAT>;<U3134> % HANGUL LETTER NIEUN
<UFFA4> <S1102>;<BASE>;<NARROW>;<UFFA4> % HALFWIDTH HANGUL LETTER NIEUN
                                % <VRNT1>    HANGUL CHOSEONG HYEOPATAKSORI NIEUN
                                % <VRNT1>    HANGUL JONGSEONG HYEOPATAKSORI NIEUN
<U3201> "<S1102><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3201> % PARENTHESIZED HANGUL NIEUN
<U3261> "<S1102><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3261> % CIRCLED HANGUL NIEUN
<U320F> "<S1102><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U320F> % PARENTHESIZED HANGUL NIEUN A
<U326F> "<S1102><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U326F> % CIRCLED HANGUL NIEUN A
<U1113> "<S1102><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U1113> % HANGUL CHOSEONG NIEUN-KIYEOK
<U11C5> "<S1102><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U11C5> % HANGUL JONGSEONG NIEUN-KIYEOK
<U1114> "<S1102><S1102>";"<BASE><BASE>";"<MIN><MIN>";<U1114> % HANGUL CHOSEONG SSANGNIEUN
<U3165> "<S1102><S1102>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3165> % HANGUL LETTER SSANGNIEUN
<U1115> "<S1102><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U1115> % HANGUL CHOSEONG NIEUN-TIKEUT
<U11C6> "<S1102><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U11C6> % HANGUL JONGSEONG NIEUN-TIKEUT
<U3166> "<S1102><S1103>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3166> % HANGUL LETTER NIEUN-TIKEUT
<U1116> "<S1102><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U1116> % HANGUL CHOSEONG NIEUN-PIEUP
<U11C7> "<S1102><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U11C7> % HANGUL JONGSEONG NIEUN-SIOS
<U3167> "<S1102><S1109>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3167> % HANGUL LETTER NIEUN-SIOS
<U11C8> "<S1102><S1140>";"<BASE><BASE>";"<MIN><MIN>";<U11C8> % HANGUL JONGSEONG NIEUN-PANSIOS
<U3168> "<S1102><S1140>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3168> % HANGUL LETTER NIEUN-PANSIOS
<U11AC> "<S1102><S110C>";"<BASE><BASE>";"<MIN><MIN>";<U11AC> % HANGUL JONGSEONG NIEUN-CIEUC
<U3135> "<S1102><S110C>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3135> % HANGUL LETTER NIEUN-CIEUC
<UFFA5> "<S1102><S110C>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFA5> % HALFWIDTH HANGUL LETTER NIEUN-CIEUC
<U11C9> "<S1102><S1110>";"<BASE><BASE>";"<MIN><MIN>";<U11C9> % HANGUL JONGSEONG NIEUN-THIEUTH
<U11AD> "<S1102><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U11AD> % HANGUL JONGSEONG NIEUN-HIEUH
<U3136> "<S1102><S1112>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3136> % HANGUL LETTER NIEUN-HIEUH
<UFFA6> "<S1102><S1112>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFA6> % HALFWIDTH HANGUL LETTER NIEUN-HIEUH

<U1103> <S1103>;<BASE>;<MIN>;<U1103> % HANGUL CHOSEONG TIKEUT (TIEUK)
<U11AE> <S1103>;<BASE>;<MIN>;<U11AE> % HANGUL JONGSEONG TIKEUT (TIEUT)
<U3137> <S1103>;<BASE>;<COMPAT>;<U3137> % HANGUL LETTER TIKEUT
<UFFA7> <S1103>;<BASE>;<NARROW>;<UFFA7> % HALFWIDTH HANGUL LETTER TIKEUT
                                % <VRNT1>    HANGUL CHOSEONG HYEOPATAKSORI TIKEUT
                                % <VRNT1>    HANGUL JONGSEONG HYEOPATAKSORI TIKEUT
<U3202> "<S1103><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3202> % PARENTHESIZED HANGUL TIKEUT
<U3262> "<S1103><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3262> % CIRCLED HANGUL TIKEUT
<U3210> "<S1103><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3210> % PARENTHESIZED HANGUL TIKEUT A
```

```
<U3270> "<S1103><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3270> % CIRCLED HANGUL TIKEUT A
<U1117> "<S1103><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U1117> % HANGUL CHOSEONG TIKEUT-KIYEOK
<U11CA> "<S1103><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U11CA> % HANGUL JONGSEONG TIKEUT-KIYEOK
<U1104> "<S1103><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U1104> % HANGUL CHOSEONG SSANGTIKEUT
<U3138> "<S1103><S1103>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3138> % HANGUL LETTER SSANGTIKEUT
<UFFA8> "<S1103><S1103>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFA8> % HALFWIDTH HANGUL LETTER SSANGTIKEUT
<U11CB> "<S1103><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U11CB> % HANGUL JONGSEONG TIKEUT-RIEUL

<U1105> <S1105>;<BASE>;<MIN>;<U1105> % HANGUL CHOSEONG RIEUL
<U11AF> <S1105>;<BASE>;<MIN>;<U11AF> % HANGUL JONGSEONG RIEUL
<U3139> <S1105>;<BASE>;<COMPAT>;<U3139> % HANGUL LETTER RIEUL
<UFFA9> <S1105>;<BASE>;<NARROW>;<UFFA9> % HALFWIDTH HANGUL LETTER RIEUL
<U3203> "<S1105><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3203> % PARENTHESIZED HANGUL RIEUL
<U3263> "<S1105><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3263> % CIRCLED HANGUL RIEUL
<U3211> "<S1105><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3211> % PARENTHESIZED HANGUL RIEUL A
<U3271> "<S1105><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3271> % CIRCLED HANGUL RIEUL A
<U11B0> "<S1105><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U11B0> % HANGUL JONGSEONG RIEUL-KIYEOK
<U313A> "<S1105><S1100>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U313A> % HANGUL LETTER RIEUL-KIYEOK
<UFFAA> "<S1105><S1100>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFAA> % HALFWIDTH HANGUL LETTER RIEUL-KIYEOK
<U11CC> "<S1105><S1100><S1109>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11CC> % HANGUL JONGSEONG RIEUL-KIYEOK-SIOS
<U3169> "<S1105><S1100><S1109>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3169> % HANGUL LETTER RIEUL-KIYEOK-SIOS
<U1118> "<S1105><S1102>";"<BASE><BASE>";"<MIN><MIN>";<U1118> % HANGUL CHOSEONG RIEUL-NIEUN
<U11CD> "<S1105><S1102>";"<BASE><BASE>";"<MIN><MIN>";<U11CD> % HANGUL JONGSEONG RIEUL-NIEUN
<U11CE> "<S1105><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U11CE> % HANGUL JONGSEONG RIEUL-TIKEUT
<U316A> "<S1105><S1103>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U316A> % HANGUL LETTER RIEUL-TIKEUT
<U11CF> "<S1105><S1103><S1112>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11CF> % HANGUL JONGSEONG RIEUL-TIKEUT-HIEUH
<U1119> "<S1105><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U1119> % HANGUL CHOSEONG SSANGRIEUL
<U11D0> "<S1105><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U11D0> % HANGUL JONGSEONG SSANGRIEUL
<U11B1> "<S1105><S1106>";"<BASE><BASE>";"<MIN><MIN>";<U11B1> % HANGUL JONGSEONG RIEUL-MIEUM
<U313B> "<S1105><S1106>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U313B> % HANGUL LETTER RIEUL-MIEUM
<UFFAB> "<S1105><S1106>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFAB> % HALFWIDTH HANGUL LETTER RIEUL-MIEUM
<U11D1> "<S1105><S1106><S1100>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11D1> % HANGUL JONGSEONG RIEUL-MIEUM-KIYEOK
<U11D2> "<S1105><S1106><S1109>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11D2> % HANGUL JONGSEONG RIEUL-MIEUM-SIOS
<U11B2> "<S1105><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U11B2> % HANGUL JONGSEONG RIEUL-PIEUP
<U313C> "<S1105><S1107>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U313C> % HANGUL LETTER RIEUL-PIEUP
<UFFAC> "<S1105><S1107>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFAC> % HALFWIDTH HANGUL LETTER RIEUL-PIEUP
<U11D3> "<S1105><S1107><S1109>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11D3> % HANGUL JONGSEONG RIEUL-PIEUP-SIOS
<U316B> "<S1105><S1107><S1109>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U316B> % HANGUL LETTER RIEUL-PIEUP-SIOS
<U11D5> "<S1105><S1107><S110B>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11D5> % HANGUL JONGSEONG RIEUL-KAPYEOUNPIEUP (RIEUL-PIEUP-
           IEUNG) % Sometimes ordered after RIEUL-PIEUP-HIEUH.
<U11D4> "<S1105><S1107><S1112>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11D4> % HANGUL JONGSEONG RIEUL-PIEUP-HIEUH
<U11B3> "<S1105><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U11B3> % HANGUL JONGSEONG RIEUL-SIOS
<U313D> "<S1105><S1109>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U313D> % HANGUL LETTER RIEUL-SIOS
<UFFAD> "<S1105><S1109>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFAD> % HALFWIDTH HANGUL LETTER RIEUL-SIOS
<U11D6> "<S1105><S1109><S1109>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11D6> % HANGUL JONGSEONG RIEUL-SSANGSIOS
<U11D7> "<S1105><S1140>";"<BASE><BASE>";"<MIN><MIN>";<U11D7> % HANGUL JONGSEONG RIEUL-PANSIOS
<U316C> "<S1105><S1140>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U316C> % HANGUL LETTER RIEUL-PANSIOS
<U111B> "<S1105><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U111B> % HANGUL CHOSEONG KAPYEOUNRIEUL (RIEUL-IEUNG) % Sometimes ordered after
           RIEUL-HIEUH.
<U11D8> "<S1105><S110F>";"<BASE><BASE>";"<MIN><MIN>";<U11D8> % HANGUL JONGSEONG RIEUL-KHIEUKH
<U11B4> "<S1105><S1110>";"<BASE><BASE>";"<MIN><MIN>";<U11B4> % HANGUL JONGSEONG RIEUL-THIEUTH
```

```
<U313E> "<S1105><S1110>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U313E> % HANGUL LETTER RIEUL-THIEUTH
<UFFAE> "<S1105><S1110>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFAE> % HALFWIDTH HANGUL LETTER RIEUL-THIEUTH
<U11B5> "<S1105><S1111>";"<BASE><BASE>";"<MIN><MIN>";<U11B5> % HANGUL JONGSEONG RIEUL-PHIEUPH
<U313F> "<S1105><S1111>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U313F> % HANGUL LETTER RIEUL-PHIEUPH
<UFFAF> "<S1105><S1111>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFAF> % HALFWIDTH HANGUL LETTER RIEUL-PHIEUPH
<U111A> "<S1105><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U111A> % HANGUL CHOSEONG RIEUL-HIEUH
<U11B6> "<S1105><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U11B6> % HANGUL JONGSEONG RIEUL-HIEUH
<U3140> "<S1105><S1112>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3140> % HANGUL LETTER RIEUL-HIEUH
<UFFB0> "<S1105><S1112>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFB0> % HALFWIDTH HANGUL LETTER RIEUL-HIEUH
<U11D9> "<S1105><S1159>";"<BASE><BASE>";"<MIN><MIN>";<U11D9> % HANGUL JONGSEONG RIEUL-YEORINHIEUH
<U316D> "<S1105><S1159>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U316D> % HANGUL LETTER RIEUL-YEORINHIEUH


<U1106> <S1106>;<BASE>;<MIN>;<U1106> % HANGUL CHOSEONG MIEUM
<U11B7> <S1106>;<BASE>;<MIN>;<U11B7> % HANGUL JONGSEONG MIEUM
<U3141> <S1106>;<BASE>;<COMPAT>;<U3141> % HANGUL LETTER MIEUM
<UFFB1> <S1106>;<BASE>;<NARROW>;<UFFB1> % HALFWIDTH HANGUL LETTER MIEUM
                              % <VRNT1>   HANGUL CHOSEONG MAREUMMIEUM
                              % <VRNT1>   HANGUL JONGSEONG MAREUMMIEUM
<U3204> "<S1106><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3204> % PARENTHESIZED HANGUL MIEUM
<U3264> "<S1106><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3264> % CIRCLED HANGUL MIEUM
<U3212> "<S1106><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3212> % PARENTHESIZED HANGUL MIEUM A
<U3272> "<S1106><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3272> % CIRCLED HANGUL MIEUM A
<U111C> "<S1106><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U111C> % HANGUL CHOSEONG MIEUM-PIEUP
<U11DA> "<S1106><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U11DA> % HANGUL JONGSEONG MIEUM-KIYEOK
<U11DB> "<S1106><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U11DB> % HANGUL JONGSEONG MIEUM-RIEUL
<U11DC> "<S1106><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U11DC> % HANGUL JONGSEONG MIEUM-PIEUP
<U316E> "<S1106><S1107>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U316E> % HANGUL LETTER MIEUM-PIEUP
<U11DD> "<S1106><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U11DD> % HANGUL JONGSEONG MIEUM-SIOS
<U316F> "<S1106><S1109>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U316F> % HANGUL LETTER MIEUM-SIOS
<U11DE> "<S1106><S1109><S1109>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11DE> % HANGUL JONGSEONG MIEUM-SSANGSIOS
<U11DF> "<S1106><S1140>";"<BASE>";"<MIN><MIN>";<U11DF> % HANGUL JONGSEONG MIEUM-PANSIOS
<U3170> "<S1106><S1140>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3170> % HANGUL LETTER MIEUM-PANSIOS
<U111D> "<S1106><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U111D> % HANGUL CHOSEONG KAPYEOUNMIEUM (MIEUM-IEUNG) % Sometimes ordered after
         MIEUM-HIEUH.
<U11E2> "<S1106><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U11E2> % HANGUL JONGSEONG KAPYEOUNMIEUM (MIEUM-IEUNG) % Sometimes ordered after
         MIEUM-HIEUH.
<U3171> "<S1106><S110B>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3171> % HANGUL LETTER KAPYEOUNMIEUM
<U11E0> "<S1106><S110E>";"<BASE><BASE>";"<MIN><MIN>";<U11E0> % HANGUL JONGSEONG MIEUM-CHIEUCH
<U11E1> "<S1106><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U11E1> % HANGUL JONGSEONG MIEUM-HIEUH


<U1107> <S1107>;<BASE>;<MIN>;<U1107> % HANGUL CHOSEONG PIEUP
<U11B8> <S1107>;<BASE>;<MIN>;<U11B8> % HANGUL JONGSEONG PIEUP
<U3142> <S1107>;<BASE>;<COMPAT>;<U3142> % HANGUL LETTER PIEUP
<UFFB2> <S1107>;<BASE>;<NARROW>;<UFFB2> % HALFWIDTH HANGUL LETTER PIEUP
<U3205> "<S1107><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3205> % PARENTHESIZED HANGUL PIEUP
<U3265> "<S1107><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3265> % CIRCLED HANGUL PIEUP
<U3213> "<S1107><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3213> % PARENTHESIZED HANGUL PIEUP A
<U3273> "<S1107><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3273> % CIRCLED HANGUL PIEUP A
<U111E> "<S1107><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U111E> % HANGUL CHOSEONG PIEUP-KIYEOK
<U3172> "<S1107><S1100>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3172> % HANGUL LETTER PIEUP-KIYEOK
<U111F> "<S1107><S1102>";"<BASE><BASE>";"<MIN><MIN>";<U111F> % HANGUL CHOSEONG PIEUP-NIEUN
```

```
<U1120> "<S1107><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U1120> % HANGUL CHOSEONG PIEUP-TIKEUT
<U3173> "<S1107><S1103>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3173> % HANGUL LETTER PIEUP-TIKEUT
<U11E3> "<S1107><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U11E3> % HANGUL JONGSEONG PIEUP-RIEUL
<U1108> "<S1107><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U1108> % HANGUL CHOSEONG SSANGPIEUP
<U3143> "<S1107><S1107>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3143> % HANGUL LETTER SSANGPIEUP
<UFFB3> "<S1107><S1107>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFB3> % HALFWIDTH HANGUL LETTER SSANGPIEUP
<U112C> "<S1107><S1107><S110B>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U112C> % HANGUL CHOSEONG KAPYEOUNSSANGPIEUP (SSANGPIEUP-IEUNG)
        % Sometimes ordered after KAPYEOUNPIEUP.
<U3179> "<S1107><S1107><S110B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3179> % HANGUL LETTER KAPYEOUNSSANGPIEUP
<U1121> "<S1107><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U1121> % HANGUL CHOSEONG PIEUP-SIOS
<U11B9> "<S1107><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U11B9> % HANGUL JONGSEONG PIEUP-SIOS
<U3144> "<S1107><S1109>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3144> % HANGUL LETTER PIEUP-SIOS
<UFFB4> "<S1107><S1109>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFB4> % HALFWIDTH HANGUL LETTER PIEUP-SIOS
<U1122> "<S1107><S1109><S1100>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1122> % HANGUL CHOSEONG PIEUP-SIOS-KIYEOK
<U3174> "<S1107><S1109><S1100>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3174> % HANGUL LETTER PIEUP-SIOS-KIYEOK
<U1123> "<S1107><S1109><S1103>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1123> % HANGUL CHOSEONG PIEUP-SIOS-TIKEUT
<U3175> "<S1107><S1109><S1103>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3175> % HANGUL LETTER 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
<U112B> "<S1107><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U112B> % HANGUL CHOSEONG KAPYEOUNPIEUP (PIEUP-IEUNG) % Sometimes ordered after
        PIEUP-HIEUH.
<U11E6> "<S1107><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U11E6> % HANGUL JONGSEONG KAPYEOUNPIEUP % Sometimes ordered after PIEUP-HIEUH.
<U3178> "<S1107><S110B>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3178> % HANGUL LETTER KAPYEOUNPIEUP
<U1127> "<S1107><S110C>";"<BASE><BASE>";"<MIN><MIN>";<U1127> % HANGUL CHOSEONG PIEUP-CIEUC
<U3176> "<S1107><S110C>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3176> % HANGUL LETTER PIEUP-CIEUC
<U1128> "<S1107><S110E>";"<BASE><BASE>";"<MIN><MIN>";<U1128> % HANGUL CHOSEONG PIEUP-CHIEUCH
<U1129> "<S1107><S1110>";"<BASE><BASE>";"<MIN><MIN>";<U1129> % HANGUL CHOSEONG PIEUP-THIEUTH
<U3177> "<S1107><S1110>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3177> % HANGUL LETTER PIEUP-THIEUTH
<U11E4> "<S1107><S1111>";"<BASE><BASE>";"<MIN><MIN>";<U11E4> % HANGUL JONGSEONG PIEUP-PHIEUPH
<U112A> "<S1107><S1111>";"<BASE><BASE>";"<MIN><MIN>";<U112A> % HANGUL CHOSEONG PIEUP-PHIEUPH
<U11E5> "<S1107><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U11E5> % HANGUL JONGSEONG PIEUP-HIEUH

<U1109> <S1109>;<BASE>;<MIN>;<U1109> % HANGUL CHOSEONG SIOS (SIEUS)
<U11BA> <S1109>;<BASE>;<MIN>;<U11BA> % HANGUL JONGSEONG SIOS (SIEUS)
<U3145> <S1109>;<BASE>;<COMPAT>;<U3145> % HANGUL LETTER SIOS
<UFFB5> <S1109>;<BASE>;<NARROW>;<UFFB5> % HALFWIDTH HANGUL LETTER SIOS
<U113C> <S1109>;"<BASE><VRNT1>";"<MIN><MIN>";<U113C> % HANGUL CHOSEONG CHITUEUMSIOS (IMEORISORISIEUS)
                                        % <VRNT1>  HANGUL JONGSEONG CHITUEUMSIOS
<U113E> <S1109>;"<BASE><VRNT2>";"<MIN><MIN>";<U113E> % HANGUL CHOSEONG CEONGCHIEUMSIOS (IMOMSORISIEUS)
                                        % <VRNT2>  HANGUL JONGSEONG CEONGCHIEUMSIOS
<U3206> "<S1109><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3206> % PARENTHESIZED HANGUL SIOS
<U3266> "<S1109><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3266> % CIRCLED HANGUL SIOS
<U3214> "<S1109><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3214> % PARENTHESIZED HANGUL SIOS A
<U3274> "<S1109><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3274> % CIRCLED HANGUL SIOS A
<U112D> "<S1109><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U112D> % HANGUL CHOSEONG SIOS-KIYEOK
<U11E7> "<S1109><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U11E7> % HANGUL JONGSEONG SIOS-KIYEOK
<U317A> "<S1109><S1100>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U317A> % HANGUL LETTER SIOS-KIYEOK
<U112E> "<S1109><S1102>";"<BASE><BASE>";"<MIN><MIN>";<U112E> % HANGUL CHOSEONG SIOS-NIEUN
<U317B> "<S1109><S1102>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U317B> % HANGUL LETTER SIOS-NIEUN
<U112F> "<S1109><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U112F> % HANGUL CHOSEONG SIOS-TIKEUT
```

```
<U11E8> "<S1109><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U11E8> % HANGUL JONGSEONG SIOS-TIKEUT
<U317C> "<S1109><S1103>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U317C> % HANGUL LETTER SIOS-TIKEUT
<U1130> "<S1109><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U1130> % HANGUL CHOSEONG SIOS-RIEUL
<U11E9> "<S1109><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U11E9> % HANGUL JONGSEONG SIOS-RIEUL
<U1131> "<S1109><S1106>";"<BASE><BASE>";"<MIN><MIN>";<U1131> % HANGUL CHOSEONG SIOS-MIEUM
<U1132> "<S1109><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U1132> % HANGUL CHOSEONG SIOS-PIEUP
<U11EA> "<S1109><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U11EA> % HANGUL JONGSEONG SIOS-PIEUP
<U317D> "<S1109><S1107>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U317D> % HANGUL LETTER SIOS-PIEUP
<U1133> "<S1109><S1107><S1100>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1133> % HANGUL CHOSEONG SIOS-PIEUP-KIYEOK
<U110A> "<S1109><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U110A> % HANGUL CHOSEONG SSANGSIOS
<U11BB> "<S1109><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U11BB> % HANGUL JONGSEONG SSANGSIOS
<U3146> "<S1109><S1109>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3146> % HANGUL LETTER SSANGSIOS
<UFFB6> "<S1109><S1109>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFB6> % HALFWIDTH HANGUL LETTER SSANGSIOS
<U113D> "<S1109><S1109>";"<BASE><VRNT1><BASE><VRNT1>";"<MIN><MIN><MIN><MIN>";<U113D> % HANGUL CHOSEONG CHITUEUMSSANGSIOS
<U113F> "<S1109><S1109>";"<BASE><VRNT2><BASE><VRNT2>";"<MIN><MIN><MIN><MIN>";<U113F> % HANGUL CHOSEONG CEONGCHIEUMSSANGSIOS
<U1134> "<S1109><S1109><S1109>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1134> % HANGUL CHOSEONG SIOS-SSANGSIOS
<U1135> "<S1109><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U1135> % HANGUL CHOSEONG SIOS-IEUNG
<U1136> "<S1109><S110C>";"<BASE><BASE>";"<MIN><MIN>";<U1136> % HANGUL CHOSEONG SIOS-CIEUC
<U317E> "<S1109><S110C>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U317E> % HANGUL LETTER SIOS-CIEUC
<U1137> "<S1109><S110E>";"<BASE><BASE>";"<MIN><MIN>";<U1137> % HANGUL CHOSEONG SIOS-CHIEUCH
<U1138> "<S1109><S110F>";"<BASE><BASE>";"<MIN><MIN>";<U1138> % HANGUL CHOSEONG SIOS-KHIEUKH
<U1139> "<S1109><S1110>";"<BASE><BASE>";"<MIN><MIN>";<U1139> % HANGUL CHOSEONG SIOS-THIEUTH
<U113A> "<S1109><S1111>";"<BASE><BASE>";"<MIN><MIN>";<U113A> % HANGUL CHOSEONG SIOS-PHIEUPH
<U113B> "<S1109><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U113B> % HANGUL CHOSEONG SIOS-HIEUH

<U1140> <S1140>;<BASE>;<MIN>;<U1140> % HANGUL CHOSEONG PANSIOS (PANISORI)
<U11EB> <S1140>;<BASE>;<MIN>;<U11EB> % HANGUL JONGSEONG PANSIOS (PANISORI)
<U317F> <S1140>;<BASE>;<COMPAT>;<U317F> % HANGUL LETTER PANSIOS

<U110B> <S110B>;<BASE>;<MIN>;<U110B> % HANGUL CHOSEONG IEUNG
<U11BC> <S110B>;<BASE>;<MIN>;<U11BC> % HANGUL JONGSEONG IEUNG
<U3147> <S110B>;<BASE>;<COMPAT>;<U3147> % HANGUL LETTER IEUNG
<UFFB7> <S110B>;<BASE>;<NARROW>;<UFFB7> % HALFWIDTH HANGUL LETTER IEUNG
<U3207> "<S110B><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3207> % PARENTHESIZED HANGUL IEUNG
<U3267> "<S110B><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3267> % CIRCLED HANGUL IEUNG
<U3215> "<S110B><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3215> % PARENTHESIZED HANGUL IEUNG A
<U3275> "<S110B><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3275> % CIRCLED HANGUL IEUNG A
<U1141> "<S110B><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U1141> % HANGUL CHOSEONG IEUNG-KIYEOK
<U11EC> "<S110B><S1100>";"<BASE><BASE>";"<MIN><MIN>";<U11EC> % HANGUL JONGSEONG IEUNG-KIYEOK
<U11ED> "<S110B><S1100><S1100>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U11ED> % HANGUL JONGSEONG IEUNG-SSANGKIYEOK
<U1142> "<S110B><S1103>";"<BASE><BASE>";"<MIN><MIN>";<U1142> % HANGUL CHOSEONG IEUNG-TIKEUT
<U1143> "<S110B><S1106>";"<BASE><BASE>";"<MIN><MIN>";<U1143> % HANGUL CHOSEONG IEUNG-MIEUM
<U1144> "<S110B><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U1144> % HANGUL CHOSEONG IEUNG-PIEUP
<U1145> "<S110B><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U1145> % HANGUL CHOSEONG IEUNG-SIOS
<U1146> "<S110B><S1140>";"<BASE><BASE>";"<MIN><MIN>";<U1146> % HANGUL CHOSEONG IEUNG-PANSIOS
<U1147> "<S110B><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U1147> % HANGUL CHOSEONG SSANGIEUNG
<U11EE> "<S110B><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U11EE> % HANGUL JONGSEONG SSANGIEUNG
<U3180> "<S110B><S110B>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3180> % HANGUL LETTER SSANGIEUNG
<U1148> "<S110B><S110C>";"<BASE><BASE>";"<MIN><MIN>";<U1148> % HANGUL CHOSEONG IEUNG-CIEUC
<U1149> "<S110B><S110E>";"<BASE><BASE>";"<MIN><MIN>";<U1149> % HANGUL CHOSEONG IEUNG-CHIEUCH
<U11EF> "<S110B><S110F>";"<BASE><BASE>";"<MIN><MIN>";<U11EF> % HANGUL JONGSEONG IEUNG-KHIEUKH
```

```
<U114A> "<S110B><S1110>";"<BASE><BASE>";"<MIN><MIN>";<U114A> % HANGUL CHOSEONG IEUNG-THIEUTH
<U114B> "<S110B><S1111>";"<BASE><BASE>";"<MIN><MIN>";<U114B> % HANGUL CHOSEONG IEUNG-PHIEUPH

<U114C> <S114C>;<BASE>;<MIN>;<U114C> % HANGUL CHOSEONG YESIEUNG (EOKEUMISORI)
<U11F0> <S114C>;<BASE>;<MIN>;<U11F0> % HANGUL JONGSEONG YESIEUNG (EOKEUMISORI)
<U3181> <S114C>;<BASE>;<COMPAT>;<U3181> % HANGUL LETTER YESIEUNG
<U11F1> "<S114C><S1109>";"<BASE><BASE>";"<MIN><MIN>";<U11F1> % HANGUL JONGSEONG YESIEUNG-SIOS
<U3182> "<S114C><S1109>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3182> % HANGUL LETTER YESIEUNG-SIOS
<U11F2> "<S114C><S1140>";"<BASE><BASE>";"<MIN><MIN>";<U11F2> % HANGUL JONGSEONG YESIEUNG-PANSIOS
<U3183> "<S114C><S1140>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3183> % HANGUL LETTER YESIEUNG-PANSIOS

<U110C> <S110C>;<BASE>;<MIN>;<U110C> % HANGUL CHOSEONG CIEUC (JIEUJ)
<U11BD> <S110C>;<BASE>;<MIN>;<U11BD> % HANGUL JONGSEONG CIEUC (JIEUJ)
<U3148> <S110C>;<BASE>;<COMPAT>;<U3148> % HANGUL LETTER CIEUC
<UFFB8> <S110C>;<BASE>;<NARROW>;<UFFB8> % HALFWIDTH HANGUL LETTER CIEUC
<U114E> <S110C>;"<BASE><VRNT1>";"<MIN><MIN>";<U114E> % HANGUL CHOSEONG CHITUEUMCIEUC (IMEORISORIJIEUJ)
                                % <VRNT1>   HANGUL JONGSEONG CHITUEUMCIEUC
<U1150> <S110C>;"<BASE><VRNT2>";"<MIN><MIN>";<U1150> % HANGUL CHOSEONG CEONGCHIEUMCIEUC (IMOMSORIJIEUJ)
                                % <VRNT2>   HANGUL JONGSEONG CEONGCHIEUMCIEUC
<U3208> "<S110C><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3208> % PARENTHESIZED HANGUL CIEUC
<U3268> "<S110C><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3268> % CIRCLED HANGUL CIEUC
<U3216> "<S110C><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3216> % PARENTHESIZED HANGUL CIEUC A
<U3276> "<S110C><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3276> % CIRCLED HANGUL CIEUC A
<U321C> "<S110C><B><S116E><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U321C> % PARENTHESIZED HANGUL CIEUC U
<U114D> "<S110C><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U114D> % HANGUL CHOSEONG CIEUC-IEUNG
<U110D> "<S110C><S110C>";"<BASE><BASE>";"<MIN><MIN>";<U110D> % HANGUL CHOSEONG SSANGCIEUC
<U3149> "<S110C><S110C>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3149> % HANGUL LETTER SSANGCIEUC
<UFFB9> "<S110C><S110C>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFB9> % HALFWIDTH HANGUL LETTER SSANGCIEUC
<U114F> "<S110C><S110C>";"<BASE><VRNT1><BASE><VRNT1>";"<MIN><MIN><MIN><MIN>";<U114F> % HANGUL CHOSEONG CHITUEUMSSANGCIEUC
<U1151> "<S110C><S110C>";"<BASE><VRNT2><BASE><VRNT2>";"<MIN><MIN><MIN><MIN>";<U1151> % HANGUL CHOSEONG CEONGCHIEUMSSANGCIEUC

<U110E> <S110E>;<BASE>;<MIN>;<U110E> % HANGUL CHOSEONG CHIEUCH
<U11BE> <S110E>;<BASE>;<MIN>;<U11BE> % HANGUL JONGSEONG CHIEUCH (JHIEUJH)
<U314A> <S110E>;<BASE>;<COMPAT>;<U314A> % HANGUL LETTER CHIEUCH
<UFFBA> <S110E>;<BASE>;<NARROW>;<UFFBA> % HALFWIDTH HANGUL LETTER CHIEUCH
<U1154> <S110E>;"<BASE><VRNT1>";"<MIN><MIN>";<U1154> % HANGUL CHOSEONG CHITUEUMCHIEUCH (IMEORISORICHIEUCH)
                                % <VRNT1>   HANGUL JONGSEONG CHITUEUMCHIEUCH
<U1155> <S110E>;"<BASE><VRNT2>";"<MIN><MIN>";<U1155> % HANGUL CHOSEONG CEONGCHIEUMCHIEUCH (IMOMSORICHIEUCH)
                                % <VRNT2>   HANGUL JONGSEONG CEONGCHIEUMCHIEUCH
<U3209> "<S110E><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3209> % PARENTHESIZED HANGUL CHIEUCH
<U3269> "<S110E><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3269> % CIRCLED HANGUL CHIEUCH
<U3217> "<S110E><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3217> % PARENTHESIZED HANGUL CHIEUCH A
<U3277> "<S110E><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3277> % CIRCLED HANGUL CHIEUCH A
<U1152> "<S110E><S110F>";"<BASE><BASE>";"<MIN><MIN>";<U1152> % HANGUL CHOSEONG CHIEUCH-KHIEUKH
<U1153> "<S110E><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U1153> % HANGUL CHOSEONG CHIEUCH-HIEUH

<U110F> <S110F>;<BASE>;<MIN>;<U110F> % HANGUL CHOSEONG KHIEUKH
<U11BF> <S110F>;<BASE>;<MIN>;<U11BF> % HANGUL JONGSEONG KHIEUKH
<U314B> <S110F>;<BASE>;<COMPAT>;<U314B> % HANGUL LETTER KHIEUKH
<UFFBB> <S110F>;<BASE>;<NARROW>;<UFFBB> % HALFWIDTH HANGUL LETTER KHIEUKH
<U320A> "<S110F><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U320A> % PARENTHESIZED HANGUL KHIEUKH
```

```
<U326A> "<S110F><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U326A> % CIRCLED HANGUL KHIEUKH
<U3218> "<S110F><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3218> % PARENTHESIZED HANGUL KHIEUKH A
<U3278> "<S110F><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3278> % CIRCLED HANGUL KHIEUKH A


<U1110> <S1110>;<BASE>;<MIN>;<U1110> % HANGUL CHOSEONG THIEUTH
<U11C0> <S1110>;<BASE>;<MIN>;<U11C0> % HANGUL JONGSEONG THIEUTH
<U314C> <S1110>;<BASE>;<COMPAT>;<U314C> % HANGUL LETTER THIEUTH
<UFFBC> <S1110>;<BASE>;<NARROW>;<UFFBC> % HALFWIDTH HANGUL LETTER THIEUTH
                            % <VRNT1>    HANGUL CHOSEONG HYEOPATAKSORI THIEUTH
                            % <VRNT1>    HANGUL JONGSEONG HYEOPATAKSORI THIEUTH
<U320B> "<S1110><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U320B> % PARENTHESIZED HANGUL THIEUTH
<U326B> "<S1110><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U326B> % CIRCLED HANGUL THIEUTH
<U3219> "<S1110><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3219> % PARENTHESIZED HANGUL THIEUTH A
<U3279> "<S1110><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U3279> % CIRCLED HANGUL THIEUTH A


<U1111> <S1111>;<BASE>;<MIN>;<U1111> % HANGUL CHOSEONG PHIEUPH
<U11C1> <S1111>;<BASE>;<MIN>;<U11C1> % HANGUL JONGSEONG PHIEUPH
<U314D> <S1111>;<BASE>;<COMPAT>;<U314D> % HANGUL LETTER PHIEUPH
<UFFBD> <S1111>;<BASE>;<NARROW>;<UFFBD> % HALFWIDTH HANGUL LETTER PHIEUPH
<U320C> "<S1111><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U320C> % PARENTHESIZED HANGUL PHIEUPH
<U326C> "<S1111><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U326C> % CIRCLED HANGUL PHIEUPH
<U321A> "<S1111><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U321A> % PARENTHESIZED HANGUL PHIEUPH A
<U327A> "<S1111><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U327A> % CIRCLED HANGUL PHIEUPH A
<U1156> "<S1111><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U1156> % HANGUL CHOSEONG PHIEUPH-PIEUP
<U11F3> "<S1111><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U11F3> % HANGUL JONGSEONG PHIEUPH-PIEUP
<U1157> "<S1111><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U1157> % HANGUL CHOSEONG KAPYEOUNPHIEUPH (PHIEUPH-IEUNG)
<U11F4> "<S1111><S110B>";"<BASE><BASE>";"<MIN><MIN>";<U11F4> % HANGUL JONGSEONG KAPYEOUNPHIEUPH (PHIEUPH-IEUNG)
<U3184> "<S1111><S110B>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3184> % HANGUL LETTER KAPYEOUNPHIEUPH


<U1112> <S1112>;<BASE>;<MIN>;<U1112> % HANGUL CHOSEONG HIEUH
<U11C2> <S1112>;<BASE>;<MIN>;<U11C2> % HANGUL JONGSEONG HIEUH
<U314E> <S1112>;<BASE>;<COMPAT>;<U314E> % HANGUL LETTER HIEUH
<UFFBE> <S1112>;<BASE>;<NARROW>;<UFFBE> % HALFWIDTH HANGUL LETTER HIEUH
<U320D> "<S1112><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U320D> % PARENTHESIZED HANGUL HIEUH
<U326D> "<S1112><B><B>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U326D> % CIRCLED HANGUL HIEUH
<U321B> "<S1112><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U321B> % PARENTHESIZED HANGUL HIEUH A
<U327B> "<S1112><B><S1161><B>";"<BASE><BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT><COMPAT>";<U327B> % CIRCLED HANGUL HIEUH A
<U11F5> "<S1112><S1102>";"<BASE><BASE>";"<MIN><MIN>";<U11F5> % HANGUL JONGSEONG HIEUH-NIEUN
<U11F6> "<S1112><S1105>";"<BASE><BASE>";"<MIN><MIN>";<U11F6> % HANGUL JONGSEONG HIEUH-RIEUL
<U11F7> "<S1112><S1106>";"<BASE><BASE>";"<MIN><MIN>";<U11F7> % HANGUL JONGSEONG HIEUH-MIEUM
<U11F8> "<S1112><S1107>";"<BASE><BASE>";"<MIN><MIN>";<U11F8> % HANGUL JONGSEONG HIEUH-PIEUP
<U1158> "<S1112><S1112>";"<BASE><BASE>";"<MIN><MIN>";<U1158> % HANGUL CHOSEONG SSANGHIEUH
<U3185> "<S1112><S1112>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3185> % HANGUL LETTER SSANGHIEUH


<U1159> <S1159>;<BASE>;<MIN>;<U1159> % HANGUL CHOSEONG YEORINHIEUH (MOKKUMEONGTHEOCHIMSORI)
<U11F9> <S1159>;<BASE>;<MIN>;<U11F9> % HANGUL JONGSEONG YEORINHIEUH (MOKKUMEONGTHEOCHIMSORI)
<U3186> <S1159>;<BASE>;<COMPAT>;<U3186> % HANGUL LETTER YEORINHIEUH


<UFDD2> <C2>;<BASE>;<MIN>;<UFDD4>      % FDD2, non-character that may be introduced by the prehandling
<UFDD3> <C3>;<BASE>;<MIN>;<UFDD3>      % FDD3, non-character that may be introduced by the prehandling
<UFDD6> <CI>;<BASE>;<MIN>;<UFDD6>      % FDD6, non-character that may be introduced by the prehanding for leading IEUNG (and YESIEUNG?)
```

```
% The 11 Hangul vowels, with clusters and compatibility forms:


<U1161> <S1161>;<BASE>;<MIN>;<U1161> % HANGUL JUNGSEONG A
<U314F> <S1161>;<BASE>;<COMPAT>;<U314F> % HANGUL LETTER A
<UFFC2> <S1161>;<BASE>;<NARROW>;<UFFC2> % HALFWIDTH HANGUL LETTER A
<U1176> "<S1161><S1169>";"<BASE><BASE>";"<MIN><MIN>";<U1176> % HANGUL JUNGSEONG A-O
<U1177> "<S1161><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U1177> % HANGUL JUNGSEONG A-U
<U1162> "<S1161><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1162> % HANGUL JUNGSEONG AE (A-I)
<U3150> "<S1161><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3150> % HANGUL LETTER AE
<UFFC3> "<S1161><S1175>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFC3> % HALFWIDTH HANGUL LETTER AE


<U1163> <S1163>;<BASE>;<MIN>;<U1163> % HANGUL JUNGSEONG YA
<U3151> <S1163>;<BASE>;<COMPAT>;<U3151> % HANGUL LETTER YA
<UFFC4> <S1163>;<BASE>;<NARROW>;<UFFC4> % HALFWIDTH HANGUL LETTER YA
<U1178> "<S1163><S1169>";"<BASE><BASE>";"<MIN><MIN>";<U1178> % HANGUL JUNGSEONG YA-O
<U1179> "<S1163><S116D>";"<BASE><BASE>";"<MIN><MIN>";<U1179> % HANGUL JUNGSEONG YA-YO
<U1164> "<S1163><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1164> % HANGUL JUNGSEONG YAE (YA-I)
<U3152> "<S1163><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3152> % HANGUL LETTER YAE
<UFFC5> "<S1163><S1175>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFC5> % HALFWIDTH HANGUL LETTER YAE


<U1165> <S1165>;<BASE>;<MIN>;<U1165> % HANGUL JUNGSEONG EO
<U3153> <S1165>;<BASE>;<COMPAT>;<U3153> % HANGUL LETTER EO
<UFFC6> <S1165>;<BASE>;<NARROW>;<UFFC6> % HALFWIDTH HANGUL LETTER EO
<U117A> "<S1165><S1169>";"<BASE><BASE>";"<MIN><MIN>";<U117A> % HANGUL JUNGSEONG EO-O
<U117B> "<S1165><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U117B> % HANGUL JUNGSEONG EO-U
<U117C> "<S1165><S1173>";"<BASE><BASE>";"<MIN><MIN>";<U117C> % HANGUL JUNGSEONG EO-EU
<U1166> "<S1165><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1166> % HANGUL JUNGSEONG E (EO-I)
<U3154> "<S1165><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3154> % HANGUL LETTER E
<UFFC7> "<S1165><S1175>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFC7> % HALFWIDTH HANGUL LETTER E


<U1167> <S1167>;<BASE>;<MIN>;<U1167> % HANGUL JUNGSEONG YEO
<U3155> <S1167>;<BASE>;<COMPAT>;<U3155> % HANGUL LETTER YEO
<UFFCA> <S1167>;<BASE>;<NARROW>;<UFFCA> % HALFWIDTH HANGUL LETTER YEO
<U117D> "<S1167><S1169>";"<BASE><BASE>";"<MIN><MIN>";<U117D> % HANGUL JUNGSEONG YEO-O
<U117E> "<S1167><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U117E> % HANGUL JUNGSEONG YEO-U
<U1168> "<S1167><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1168> % HANGUL JUNGSEONG YE (YEO-I)
<U3156> "<S1167><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3156> % HANGUL LETTER YE
<UFFCB> "<S1167><S1175>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFCB> % HALFWIDTH HANGUL LETTER YE


<U1169> <S1169>;<BASE>;<MIN>;<U1169> % HANGUL JUNGSEONG O
<U3157> <S1169>;<BASE>;<COMPAT>;<U3157> % HANGUL LETTER O
<UFFCC> <S1169>;<BASE>;<NARROW>;<UFFCC> % HALFWIDTH HANGUL LETTER O
<U116A> "<S1169><S1161>";"<BASE><BASE>";"<MIN><MIN>";<U116A> % HANGUL JUNGSEONG WA (O-A)
```

```
<U3158> "<S1169><S1161>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3158> % HANGUL LETTER WA
<UFFCD> "<S1169><S1161>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFCD> % HALFWIDTH HANGUL LETTER WA
<U116B> "<S1169><S1161><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U116B> % HANGUL JUNGSEONG WAE (O-A-I)
<U3159> "<S1169><S1161><S1175>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3159> % HANGUL LETTER WAE
<UFFCE> "<S1169><S1161><S1175>";"<BASE><BASE><BASE>";"<NARROW><NARROW><NARROW>";<UFFCE> % HALFWIDTH HANGUL LETTER WAE
<U117F> "<S1169><S1165>";"<BASE><BASE>";"<MIN><MIN>";<U117F> % HANGUL JUNGSEONG O-EO
<U1180> "<S1169><S1165><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1180> % HANGUL JUNGSEONG O-E (O-EO-I)
<U1181> "<S1169><S1167><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1181> % HANGUL JUNGSEONG O-YE (O-YEO-I)
<U1182> "<S1169><S1169>";"<BASE><BASE>";"<MIN><MIN><MIN>";<U1182> % HANGUL JUNGSEONG O-O
<U1183> "<S1169><S116E>";"<BASE><BASE>";"<MIN><MIN><MIN>";<U1183> % HANGUL JUNGSEONG O-U
<U116C> "<S1169><S1175>";"<BASE><BASE>";"<MIN><MIN><MIN>";<U116C> % HANGUL JUNGSEONG OE (O-I)
<U315A> "<S1169><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U315A> % HANGUL LETTER OE
<UFFCF> "<S1169><S1175>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFCF> % HALFWIDTH HANGUL LETTER OE

<U116D> <S116D>;<BASE>;<MIN>;<U116D> % HANGUL JUNGSEONG YO
<U315B> <S116D>;<BASE>;<COMPAT>;<U315B> % HANGUL LETTER YO
<UFFD2> <S116D>;<BASE>;<NARROW>;<UFFD2> % HALFWIDTH HANGUL LETTER YO
<U1184> "<S116D><S1163>";"<BASE><BASE>";"<MIN><MIN>";<U1184> % HANGUL JUNGSEONG YO-YA
<U3187> "<S116D><S1163>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3187> % HANGUL LETTER YO-YA
<U1185> "<S116D><S1163><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1185> % HANGUL JUNGSEONG YO-YAE (YO-YA-I)
<U3188> "<S116D><S1163><S1175>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U3188> % HANGUL LETTER YO-YAE
<U1186> "<S116D><S1167>";"<BASE><BASE>";"<MIN><MIN>";<U1186> % HANGUL JUNGSEONG YO-YEO
<U1187> "<S116D><S1169>";"<BASE><BASE>";"<MIN><MIN>";<U1187> % HANGUL JUNGSEONG YO-O
<U1188> "<S116D><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1188> % HANGUL JUNGSEONG YO-I
<U3189> "<S116D><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3189> % HANGUL LETTER YO-I

<U116E> <S116E>;<BASE>;<MIN>;<U116E> % HANGUL JUNGSEONG U
<U315C> <S116E>;<BASE>;<COMPAT>;<U315C> % HANGUL LETTER U
<UFFD3> <S116E>;<BASE>;<NARROW>;<UFFD3> % HALFWIDTH HANGUL LETTER U
<U1189> "<S116E><S1161>";"<BASE><BASE>";"<MIN><MIN>";<U1189> % HANGUL JUNGSEONG U-A
<U118A> "<S116E><S1161><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U118A> % HANGUL JUNGSEONG U-AE (U-A-I)
<U116F> "<S116E><S1165>";"BASE><BASE>";"<MIN><MIN>";<U116F> % HANGUL JUNGSEONG WEO (U-EO)
<U315D> "<S116E><S1165>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U315D> % HANGUL LETTER WEO
<UFFD4> "<S116E><S1165>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFD4> % HALFWIDTH HANGUL LETTER WEO
<U118B> "<S116E><S1165><S1173>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U118B> % HANGUL JUNGSEONG U-EO-EU
<U1170> "<S116E><S1165><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1170> % HANGUL JUNGSEONG WE (U-EO-I)
<U315E> "<S116E><S1165><S1175>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U315E> % HANGUL LETTER WE
<UFFD5> "<S116E><S1165><S1175>";"<BASE><BASE><BASE>";"<NARROW><NARROW><NARROW>";<UFFD5> % HALFWIDTH HANGUL LETTER WE
<U118C> "<S116E><S1167><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U118C> % HANGUL JUNGSEONG U-YE
<U118D> "<S116E><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U118D> % HANGUL JUNGSEONG U-U
<U1171> "<S116E><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1171> % HANGUL JUNGSEONG WI (U-I)
<U315F> "<S116E><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U315F> % HANGUL LETTER WI
<UFFD6> "<S116E><S1175>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFD6> % HALFWIDTH HANGUL LETTER WI

<U1172> <S1172>;<BASE>;<MIN>;<U1172> % HANGUL JUNGSEONG YU
<U3160> <S1172>;<BASE>;<COMPAT>;<U3160> % HANGUL LETTER YU
<UFFD7> <S1172>0;<BASE>;<NARROW>;<UFFD7> % HALFWIDTH HANGUL LETTER YU
<U118E> "<S1172><S1161>";"<BASE><BASE>";"<MIN><MIN>";<U118E> % HANGUL JUNGSEONG YU-A
<U118F> "<S1172><S1165>";"<BASE><BASE>";"<MIN><MIN>";<U118F> % HANGUL JUNGSEONG YU-EO
<U1190> "<S1172><S1165><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1190> % HANGUL JUNGSEONG YU-E (YU-EO-I)
<U1191> "<S1172><S1167>";"<BASE><BASE>";"<MIN><MIN>";<U1191> % HANGUL JUNGSEONG YU-YEO
```

```
<U318A> "<S1172><S1167>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U318A> % HANGUL LETTER YU-YEO
<U1192> "<S1172><S1167><S1175>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1192> % HANGUL JUNGSEONG YU-YE (YU-YEO-I)
<U318B> "<S1172><S1167><S1175>";"<BASE><BASE><BASE>";"<COMPAT><COMPAT><COMPAT>";<U318B> % HANGUL LETTER YU-YE
<U1193> "<S1172><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U1193> % HANGUL JUNGSEONG YU-U
<U1194> "<S1172><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1194> % HANGUL JUNGSEONG YU-I
<U318C> "<S1172><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U318C> % HANGUL LETTER YU-I

<U1173> <S1173>;<BASE>;<MIN>;<U1173> % HANGUL JUNGSEONG EU
<U3161> <S1173>;<BASE>;<COMPAT>;<U3161> % HANGUL LETTER EU
<UFFDA> <S1173>;<BASE>;<NARROW>;<UFFDA> % HALFWIDTH HANGUL LETTER EU
<U1195> "<S1173><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U1195> % HANGUL JUNGSEONG EU-U
<U1196> "<S1173><S1173>";"<BASE><BASE>";"<MIN><MIN>";<U1196> % HANGUL JUNGSEONG EU-EU
<U1174> "<S1173><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U1174> % HANGUL JUNGSEONG YI (EU-I)
<U3162> "<S1173><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U3162> % HANGUL LETTER YI
<UFFDB> "<S1173><S1175>";"<BASE><BASE>";"<NARROW><NARROW>";<UFFDB> % HALFWIDTH HANGUL LETTER YI
<U1197> "<S1173><S1175><S116E>";"<BASE><BASE><BASE>";"<MIN><MIN><MIN>";<U1197> % HANGUL JUNGSEONG YI-U (EU-I-U)

<U1175> <S1175>;<BASE>;<MIN>;<U1175> % HANGUL JUNGSEONG I
<U3163> <S1175>;<BASE>;<COMPAT>;<U3163> % HANGUL LETTER I
<UFFDC> <S1175>;<BASE>;<NARROW>;<UFFDC> % HALFWIDTH HANGUL LETTER I
<U1198> "<S1175><S1161>";"<BASE><BASE>";"<MIN><MIN>";<U1198> % HANGUL JUNGSEONG I-A
<U1199> "<S1175><S1163>";"<BASE><BASE>";"<MIN><MIN>";<U1199> % HANGUL JUNGSEONG I-YA
<U119A> "<S1175><S1169>";"<BASE><BASE>";"<MIN><MIN>";<U119A> % HANGUL JUNGSEONG I-O
<U119B> "<S1175><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U119B> % HANGUL JUNGSEONG I-U
<U119C> "<S1175><S1173>";"<BASE><BASE>";"<MIN><MIN>";<U119C> % HANGUL JUNGSEONG I-EU
<U119D> "<S1175><S119E>";"<BASE><BASE>";"<MIN><MIN>";<U119D> % HANGUL JUNGSEONG I-ARAEA

<U119E> <S119E>;<BASE>;<MIN>;<U119E> % HANGUL JUNGSEONG ARAEA (ARAIA)
<U318D> <S119E>;<BASE>;<COMPAT>;<U318D> % HANGUL LETTER ARAEA
<U119F> "<S119E><S1165>";"<BASE><BASE>";"<MIN><MIN>";<U119F> % HANGUL JUNGSEONG ARAEA-EO
<U11A0> "<S119E><S116E>";"<BASE><BASE>";"<MIN><MIN>";<U11A0> % HANGUL JUNGSEONG ARAEA-U
<U11A1> "<S119E><S1175>";"<BASE><BASE>";"<MIN><MIN>";<U11A1> % HANGUL JUNGSEONG ARAEA-I
<U318E> "<S119E><S1175>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U318E> % HANGUL LETTER ARAEAE
<U11A2> "<S119E><S119E>";"<BASE><BASE>";"<MIN><MIN>";<U11A2> % HANGUL JUNGSEONG SSANGARAEA


<UFDD4> <V2>;<BASE>;<MIN>;<UFDD4>        % FDD4, non-character that may be introduced by the prehandling
<UFDD5> <V3>;<BASE>;<MIN>;<UFDD5>        % FDD5, non-character that may be introduced by the prehandling
```

%%% END SUGGESTED HANGUL EXCERPT FOR 14651 CTT

3 Tailorings

3.1 Suggested tailoring for R. of Korea ordering

The official ordering appears to be (Han'gŭl matchumpŏp, 1988):

For modern leading consonants and leading consonants clusters:

ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
k/g	kk	n	d/t	tt	l/r	m	p/b	pp	s	ss	-	ch/j	tch	ch'	k'	t'	p'	h

For modern trailing consonants and trailing consonant clusters:

ㄱ	ㄲ	ㄳ	ㄴ	ㄵ	ㄶ	ㄷ	ㄹ	ㄺ	ㄻ	ㄼ	ㄽ	ㄾ	ㄿ	ㅀ	ㅁ	ㅂ	ㅄ	ㅅ	ㅆ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
k/g	kk	ks	n	nj	nh	t/d	l/r	lk/lg	lm	lp/lb	ls	lt'	lp'	lh	m	p/b	ps	s	ss	ng	ch/j	ch'	k'	t'	p'	h

For modern vowels and vowel clusters:

ㅏ	ㅐ	ㅑ	ㅒ	ㅓ	ㅔ	ㅕ	ㅖ	ㅗ	ㅘ	ㅙ	ㅚ	ㅛ	ㅜ	ㅝ	ㅞ	ㅟ	ㅠ	ㅡ	ㅢ	ㅣ
a	ae	ya	yae	ŏ	e	yŏ	ye	o	wa	wae	oe	yo	u	wŏ	we	wi	yu	ŭ	ŭi	i

This order is a subset of the ordering that results from the above suggested CTT excerpt, without any tailoring. Note that the syllable and cluster based ordering that is used for Hangul has been done by assigning the weights as implied by the desired order as follows (X is any character that cannot be part of a Hangul syllable):

- $L_1V < L_1LV$ implies that (initial V) < (non-initial L); done via insertion of U+FDD1 after the leading consonants.
- $L_1V_1L < L_1V_1T$ implies that (initial L) < T; done via insertion of U+FDD1 after the last letter of a syllable.
- LVX < LVT implies that X < T; done via insertion of U+FDD1 after the last letter of a syllable.
- $L_1V_1T < L_1V_1V$ implies that T < (non-initial V); done by listing the Hangul (jamo) letter (not filler) weights thus.
- (non-initial L) and T either way (merged chosen).

Alternatively, a U+FDD1 could be introduced also after the vowels. This would be slightly more systematic, and lead to that vowels need not be ordered after the (trailing) consonants, but would insert more U+FDD1 characters. Note that if there is no trailing consonant, there will be two U+FDD1 at the end of the syllable in this alternative.

If desired, in order to order the *"KAPYEOUN..."* constructs, both precomposed and constructed, as if for those IEUNG were the last consonant, some tailoring is needed. However, this only results in a very minor change to the ordering, and may be regarded as an unnecessary complication.
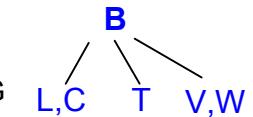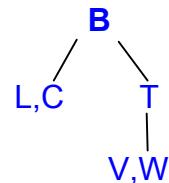
3.2 Suggested tailoring for Han'gŭl match'umpŏp t'ong'iran ordering

The Han'gŭl match'umpŏp t'ong'iran (1933) ordering appears to systematically order clusters of two after singletons, and clusters of three (or more) thereafter. The author has no information about how the "KAPYEOUN..." constructs were supposed to be ordered, nor other historic Hangul clusters. The intended ordering appears to be:

For modern leading consonants and leading consonants clusters:

ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	ㄲ	ㄸ	ㅃ	ㅆ	ㅉ
k/g	n	d/t	l/r	m	p/b	s	-	ch/j	ch'	k'	t'	p'	h	kk	tt	pp	ss	tch

For modern trailing consonants and consonants clusters:

ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	ㄲ	ㄳ	ㄵ	ㄶ	ㄺ	ㄻ	ㄼ	ㄽ	ㄾ	ㄿ	ㅀ	ㅄ	ㅆ
k/g	n	d/t	l/r	m	p/b	s	ng	ch/j	ch'	k'	t'	p'	h	kk	ks	nj	nh	lk/lg	lm	lp/lb	Ls	lt'	lp'	lh	ps	ss

For modern vowels and vowels clusters:

ㅏ	ㅑ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅠ	ㅡ	ㅣ	ㅐ	ㅒ	ㅔ	ㅖ	ㅘ	ㅚ	ㅝ	ㅟ	ㅢ	ㅙ	ㅞ
a	ya	ŏ	yŏ	o	yo	u	yu	ŭ	i	ae	yae	e	ye	wa	oe	wŏ	wi	ŭi	wae	we

This order is a subset of the ordering that results from the suggested CTT excerpt together with the following tailoring. Some, as yet unused, weights are needed for this tailoring. For convenience here, they are already present in the suggested CTT excerpt above. The change thus affects only the prehandling phase, which adds a step.

5. Insertion of cluster size characters (used just for Hangul) **C2**, **C3**, **V2**, and **V3** in the following positions in a Hangul syllable: L+ **B** V+ T* **B** M* → (**C**n)?L+ **B** (**V**n)?V+ (**C**n)?T* **B** M* and H (C+|H) **B** (W+|H) (C+|H) **B** M* → H (**C**n)?(C+|H) (**V**n)?(W+|H) (**C**n)?(C+|H) **B** M*. They are inserted in this way: **C2** is inserted in front of clusters of two consonants, **C3** is inserted in front of clusters of three or more consonants, **V2** is inserted in front of clusters of two vowels, and **V3** is inserted in front of clusters of three or more vowels.

Just like that the consonants are given one weighting independent of L-ness or T-ness there is no need to have different weightings for two-consonant cluster weighting for L and T and similarly for three-consonant cluster weighting for L and T.

```
              B
            /   \
         L,C      T
          |       |
         c2       C2
          |       |
         c3       C3
                  |
                 V,W
                  |
                 V2
                  |
                 V3
```
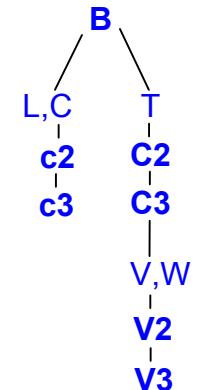
3.3 DPR of Korea ordering

The DPRK ordering may to be a modification of the Han'gŭl match'umpŏp t'ong'iran ordering. The ordering appears to be:

For leading consonants (note that IEUNG is last, maybe because it is silent when leading, where neither of the lists in SC2/WG2/N2246 put it, but how it is ordered in JTC1/N5999):

ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	ㄲ	ㄸ	ㅃ	ㅆ	ㅉ	ㅇ
k/g	n	d/t	l/r	m	p/b	s	ch/j	ch'	k'	t'	p'	h	kk	tt	pp	ss	tch	-

For trailing consonants (note the mixed approach):

ㄱ	ㄳ	ㄴ	ㄵ	ㄶ	ㄷ	ㄹ	ㄺ	ㄻ	ㄼ	ㄽ	ㄾ	ㄿ	ㅀ	ㅁ	ㅂ	ㅄ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	ㄲ	ㅆ
k/g	ks	n	nj	nh	t/d	l/r	lk/lg	lm	lp/lb	ls	lt'	lp'	lh	m	p/b	ps	s	ng	ch/j	ch'	k'	t'	p'	h	kk	ss

For vowels, either (from JTC1/N5999, and SC2WG2/N2246 page 3):

ㅏ	ㅑ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅠ	ㅡ	ㅣ	ㅐ	ㅒ	ㅔ	ㅖ	ㅚ	ㅟ	ㅢ	ㅘ	ㅝ	ㅙ	ㅞ
a	ya	ǒ	yǒ	o	yo	u	yu	ǔ	i	ae	yae	e	ye	oe	wi	ǔi	wa	wǒ	wae	we

or (from SC2/WG2/N2246, page 4):

ㅏ	ㅑ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅠ	ㅡ	ㅣ	ㅐ	ㅒ	ㅔ	ㅖ	ㅘ	ㅙ	ㅚ	ㅝ	ㅞ	ㅟ	ㅢ
a	ya	ǒ	yǒ	o	yo	u	yu	ǔ	i	ae	yae	e	ye	wa	wae	oe	wǒ	we	wi	ǔi

B
L,C   T
c2   C2
c3   C3
ci
           V,W
           V2
           V3

   This appears to be a rather mixed and confusing approach. In addition the ordering given in N5999 for historic Hangul is entirely unsupported, and also very complex to describe in a general manner (handling also single-letter Jamo sequences). E.g. the `"KAPYEOUN..."` constructs are ordered among single letters, how YESIEUNG is ordered depends on if it is leading or trailing, etc. This kind of complexity is unhelpful for users of the ordering, that may wish to make a lookup in a phonebook, dictionary, or book index with as easy rules as possible.

   The complexities that appear to be introduced in SC22/WG20/N867, when generalised, appear not to be motivated. Indeed only a small part of the complexities appear in the to ISO WGs told DPRK ordering for modern Hangul, which in turn is different from the ordering in some of the sources copied into SC2/WG2/N2246. Doing lots of work to get an unfounded ordering for historic (and modern) Hangul appears misdirected. Indeed, even the subset that orders modern Hangul in the official DPRK way has been criticised for its confusing approach, special-treating vowel clusters that end in a Hangul vowel "I", as well as the mixed approach for the trailing consonants. It is probably less confusing to use the Han'gŭl match'umpŏp t'ong'iran ordering instead, which is otherwise similar, but without the
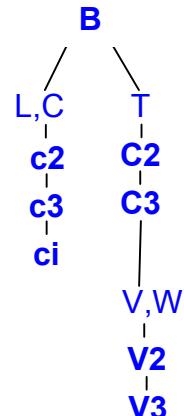
confusion, especially when historic forms of Hangul is to be handled, possibly with the only addition to order leading IEUNG as the last consonant. Or indeed, use the suggested CTT excerpt as given above without any further tailoring. The latter would have the advantage of being true to history (i.e. the original presentation of the Hangul script more than 500 years ago), and would also enable a unified ordering of Hangul for both Koreas.

4 Shortening the ordering keys

Since the prehandling introduces many new characters, the lengths of the collation keys for Hangul strings get longer than the number of original characters. For levels 2 to 4 ordinary run-length based compression of the subkey representation just result in higher values for the run-lengths. But the weights actually change at level 1, so that method cannot profitably be used on level 1. Instead it is better to introduce contractions, just to shorten the keys, not for changing the order, of the following forms; all are given single weights at each level:

```
        B
   BV,BW
   L,C       T
        V,W
```

- <U+FDD1, a Hangul jungseong/vowel>. These contractions are weighted before any leading Hangul consonant, in the same order as the Hangul vowels.

<U+FDD1, a Hangul choseong/leading consonant> can be contracted only if <U+FDD1, any letter of any script> is also contracted for all letters and all scripts (which is not realistic, would double the size of the weighting table, and there would be more contractions than there are letters).

If a tailoring that orders first on the length of clusters is used (see the given the tailorings), have also contractions of the forms:

```
             B
     BV,BW
   BV2V,BV2W
   BV3V,BV3W
     L,C        T
   c2L,c2C      c2T
   c3L,c3C      c3T
   ciL,ciC
              V,W
           V2V,V2W
           V3V,V3W
```

- <U+FDD1, U+FDD4, a Hangul jungseong/vowel>: After all <U+FDD1, a Hangul jungseong/vowel>.

- <U+FDD1, U+FDD5, a Hangul jungseong/vowel>: Thereafter.

- <U+FDD2, a Hangul choseong/leading consonant> and <U+FDD2, a Hangul jongseong/trailing consonant>: After all (single) consonants.

- <U+FDD3, a Hangul choseong/leading consonant> and <U+FDD3, a Hangul jongseong/trailing consonant>: Thereafter.

- <U+FDD6, HANGUL CHOSEONG IEUNG>: Thereafter, but still before all (single) Hangul vowels.

This will generate multiple copies of the Hangul letters in the ordering table, but not so many that linguistic considerations need be made to reduce the number, especially not if no cluster Jamos or cluster letters are used.
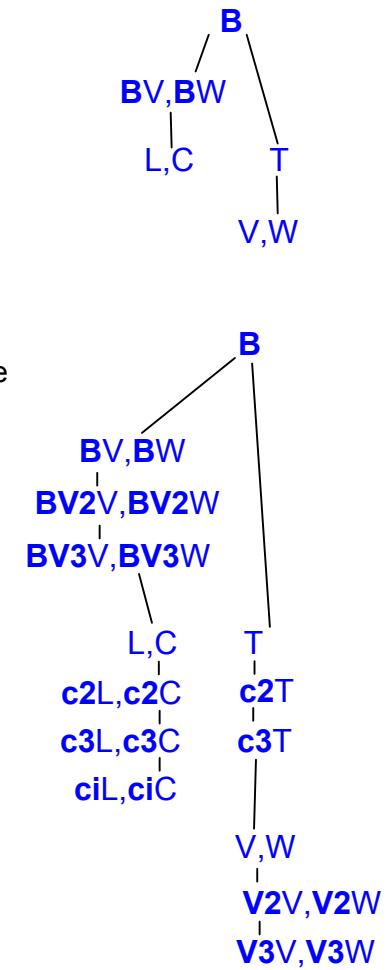
The number of weights for a Hangul string will then be nearly the same as for any other non-cluster-ordered alphabetic script: one weight for each letter in the string, but with just one additional weight per syllable.

One can remove yet another **B** character (cluster end mark, U+FDD1 used here) for the Jamo Hangul syllables, the one after the last Hangul syllable in a Hangul substring (reason: there is no ordering requirement between Xs, **B**Xs, and **BB**Xs, so they can collate identically; X is any character that cannot be part of a Hangul syllable). however, this will create an anomaly for the ordering relationship between compatibility Hangul letter strings and Jamo Hangul strings: the trailing consonant in the compatibility Hangul string will order before all the trailing Hangul Jamo consonants, and not intercollate, since this change breaks the ordering relationship between the Ts and the Cs. Since the proper handling of Hangul compatibility strings is not a priority, this may be acceptable. If so, the **B** at the end of the Jamo trailing consonants cluster can be removed, if the jongseong (the Jamo training consonants) are ordered after every other script (and the Hangul vowels still thereafter).

```
            B
           / \
   BV,BW      X
    |         |
   L,C        |
      \       T
       \      |
        \    V,W
```

Another possibility (orthogonal to the previous possibility of ordering trailing consonants Jamo after all scripts) is to make a list of "all" consonant (leading and trailing) and vowel clusters and make them into contractions with a single weight at level 1 each, in the order desired. "All" can limited to clusters of at most three Hangul letters in total, however represented, so that the list is finite. This will shorten the collation keys a bit, and will also relax the ordering conditions so that first vowels need not be ordered before the leading consonants, nor need non-first vowels come after trailing consonants. This obviates the need to have a **B** after the leading consonants. If all the **B**-removal optimisations are combined, nearly no **B** need be inserted during the prehandling except if the vowels in a syllable are all FILLERs, in which case the **B** after the leading consonants cannot be removed without introducing another small potential anomaly, that may be negligible. **To check**: can **BB**Ts collate identically with **B**Ts, like that the latter can collate identically with Ts, and can **BB**Xs collate identically with **B**Xs like that the latter can collate identically with Xs?

```
                 B
              /  |  \
      L,C,LL,   X    V,W,
      CC,LLL,   |    VV,WW,
        CCC     |    VVV,
           \    |    WWW
          T,TT,
            TTT
```

However, the last approach is unpractical: there are too many combinations (more than ten thousand, even if cluster characters are expanded). A suggestion is to make a "linguistic selection" of a few hundred combinations that are actually used. This makes this last approach practical. But the selection is both error-prone an may easily change over time. Making such selections has been acted upon repeatedly for Hangul, apparently with different(!) results: the Unicode/10646 list, Microsoft's list, and the list from DPRK, which all differ. As far as the author knows, it has not yet been acted upon for other cluster collated scripts, but if it has been or will be then chances are great that there will be multiple lists for each script. Note that clusters that are not in the list of contractions will be *mistreated* in the ordering if this approach is used. This approach should therefore ***not*** be preferred, or even ever used.
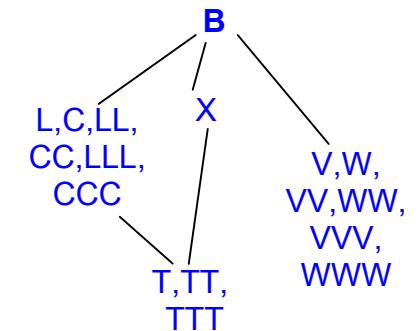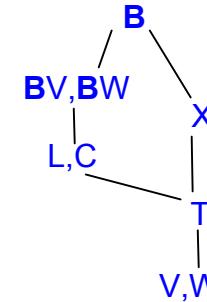
5 New Hangul characters

References to some as yet unallocated single-letter Hangul Jamo variant letters are given in comments in the ordering rules text above. If evidence for them are presented, they should be allocated. No more cluster Jamos should be allocated. Nor any more cluster compatibility Hangul letters. But separate variants, like:

```
115A HANGUL CHOSEONG HYEOPATAKSORI NIEUN (possibly missing variant)
115B HANGUL CHOSEONG HYEOPATAKSORI TIKEUT (possibly missing variant)
115C HANGUL CHOSEONG MAREUMMIEUM (possibly missing variant)
115D HANGUL CHOSEONG HYEOPATAKSORI THIEUTH (possibly missing variant)

11A4 HANGUL JONGSEONG HYEOPATAKSORI NIEUN (trailing counterpart of possibly missing variant)
11A5 HANGUL JONGSEONG HYEOPATAKSORI TIKEUT (trailing counterpart of possibly missing variant)
11A6 HANGUL JONGSEONG MAREUMMIEUM (trailing counterpart of possibly missing variant)
11A7 HANGUL JONGSEONG HYEOPATAKSORI THIEUTH (trailing counterpart of possibly missing variant)

11FA HANGUL JONGSEONG CHITUEUMSIOS (trailing counterpart of existing variant)
11FB HANGUL JONGSEONG CEONGCHIEUMSIOS (trailing counterpart of existing variant)
11FC HANGUL JONGSEONG CHITUEUMCIEUC (trailing counterpart of existing variant)
11FD HANGUL JONGSEONG CEONGCHIEUMCIEUC (trailing counterpart of existing variant)
11FE HANGUL JONGSEONG CHITUEUMCHIEUCH (trailing counterpart of existing variant)
11FF HANGUL JONGSEONG CEONGCHIEUMCHIEUCH (trailing counterpart of existing variant)
```

maybe should be allocated (maybe also as compatibility letters), subject to further evidence.

6 Other clustered collations

Also other scripts may be collated on a "cluster" basis. The most common "cluster" is a word. A common and simple way of getting a word based collation on level 1 is very similar to the technique used above for scripts that use space between words: make space level 1 significant and ordered before any letter. This assumes that there is a single space between words, or to be precise, there is a single space *after* each word, or that other punctuation is also significant at level 1. Since this cannot be relied upon to be the input, in particular there may be multiple punctuation, a better alternative is to insert a special end marker (U+FDD1 as used above, or whatever else is chosen, can be reused, but see below) just after each word, still ignoring space and other punctuation at levels 1-3. (Making a cluster contraction for each word is not a realistic approach!)

Another way of implementing word based clustering is to parse the string into words, that are each given full (all levels) keys, and then compare sequences of collation keys. This gives a stricter division of the words.

Numerals for natural numbers as well as other numbers is another common kind of logical character string cluster that should be ordered as a cluster too (see ISO/IEC 14651). These should ordered basis of value represented. One way to get that is to use something very similar to the technique used for the length marking of Hangul letter clusters in the tailoring above. Even though, like for the length separated Hangul clusters (as long as there are at most three letters in a cluster), the cluster terminator character **B** can be removed for natural and integer numerals, they need to be kept for fractional numerals (the initial 2 is a length mark for the integer part of the numeral): 2:25.63 should come after 2:25.6F (which it does not by itself since digits are lighter than letters); which one gets by inserting the **B** character at the end of the numeral cluster: 2:25.63**B** does come after 2:25.6**B**F. Note that the **B** (U+FDD1 in this paper) must be ordered before any digits. An alternative that obviates the need to insert a **B** at the end of a numeral, is to collate all digits *after* all scripts, just before the Hangul trailing consonants Jamo.

Khmer uses a similar clustered ordering as Hangul (but Khmer does not have any clusters allocated as characters). Again, this can be achieved by inserting a special end marker (U+FDD1 can be reused) for the clusters. Like for Hangul, it is possible to make a contraction for each cluster, but to keep the number of contractions down, a selection must then be made. And like for Hangul, such a selection would be error-prone, and may miss cases that are, or will be, actually used. Instead, to shorten the keys a bit, the technique explained in section 4 above can be reapplied.

Other scripts that may be ordered by some smaller-than-words clustering, can be handled in the same way: insert a cluster terminator, **B** (U+FDD1 in this paper), at the end of each cluster. Some further refinements may be needed, as demonstrated in the Hangul tailorings and the key shortening method described above.

If syllable/subsyllable clustering is desired at the same time as word clustering, two cluster end characters are needed. One for the syllables/subsyllables clustering, and one for the words clustering. The one for the word clustering must be weighted lighter than the one for the syllables/subsyllables clustering.

```
collating_symbol <W>        % FDD0, special word/numeral end character. Ordered before any letters or digits of any script.
collating_symbol <B>        % FDD1, special subword cluster end character. Ordered before any letters or digits of any script.

<W>            % FDD0, word/numeral end mark introduced by the prehandling. [RECORD SEPARATOR?]
<B>            % FDD1, subword cluster end mark introduced by the prehandling. [UNIT SEPARATOR?]

<UFDD0> <W>;<BASE>;<MIN>;<UFDD0>      % FDD0, non-character that is introduced by the prehandling
<UFDD1> <B>;<BASE>;<MIN>;<UFDD1>      % FDD1, non-character that is introduced by the prehandling
```

Etcetera for hierarchical clustering (phrases, sentences, paragraphs, ...). The lower level cluster end marks can be systematically removed just before a higher level cluster end mark, as well as just before end of string.

7 Acknowledgements

Many thanks to Jungshik Shin for providing an excerpt from SOHN Ho-Min's paper *Orthographic divergence in South and North Korea: Toward a unified spelling system*, for helping the author understand the basic ordering requirements for Hangul, and understanding the mapping from compatibility Hangul letters to Hangul Jamo. Also thanks to Mark Davis for comments that resulted in substantial changes and improvements. Any remaining errors or shortcomings are of course mine.

8 References

ISO/IEC 10646-1:2000	Information Technology – Universal multiple-octet coded character set (UCS), Part 1, second edition.
Unicode 2.0	The Unicode standard, version 2.0. (In particular the compatibility decompositions in the UCD and (algorithmic) canonical decomposition of Hangul syllables.)
Unicode 3.0	The Unicode standard, version 3.0. (In particular the (algorithmic) canonical decomposition of Hangul syllables.)
UCD 2.0	Unicode character database, version 2.0.
UCD 3.1	Unicode character database, version 3.1 (esp. new non-characters).
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 [sic.; clusters and letters] 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 N879   Kent Karlsson, Revised draft ISO/IEC 14651:2001 CTT tailoring delta for D.P.R. Korea. Review of N867.

Han'gŭl matchumpŏp (The hangŭl spelling conventions). Ministry of Education. Mun'gyobu. Seoul. 1988.

Han'gŭl match'umpŏp t'ongi'iran (A proposition for unified han'gŭl spelling conventions). Chosŏnŏ Hakhoe (Korean Language Association). Seoul. 1933. (There may possibly be different versions of this document.)

Kaejŏnghan Chosŏnmal kyubŏmjip (A revised collection of Korean language norms). Kugŏ sajŏng wiwonhoe (Korean Language Assessment Committee). Sahoegwahak Ch'ulp'ansa. P'yŏng'yang. 1988.

Orthographic divergence in South and North Korea: Toward a unified spelling system. SOHN Ho-Min. In 'The Korean Alphabet', ed. Young-Key Kim-Renaud, University of Hawaii Press. 1997.

Segmental collation and the UCA. Mark Davis. "Very draft", 2001-11-16.

——— end ———