

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title: *Proposal to encode two Kana characters concerning YE.*

2. Requester's name: *Nozomu Katō*

3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*

4. Submission date: *2008-1-14*

5. Requester's reference (if applicable):

6. Choose one of the following:

This is a complete proposal:

(or) More information will be provided later:

Yes

No

B. Technical – General

1. Choose one of the following:

a. This proposal is for a new script (set of characters):

Proposed name of script:

No

b. The proposal is for addition of character(s) to an existing block:

*Yes for one,
No for one (due
to no empty
point in the
proper block).*

Name of the existing block:

HIRAGANA (and KATAKANA)

2

2. Number of characters in proposal:

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary

B.1-Specialized (small collection)

B.2-Specialized (large collection)

C-Major extinct

D-Attested extinct

E-Minor extinct

F-Archaic Hieroglyphic or Ideographic

G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided?

a. If YES, are the names in accordance with the "character naming guidelines"

in Annex L of P&P document?

Yes

b. Are the character shapes attached in a legible form suitable for review?

Yes

5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Nozomu Katō

If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*

b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *Yes*

7. Special encoding issues:

Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *Yes*

Information about the sorting order is enclosed.

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N3102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	<input checked="" type="checkbox"/> No
If YES explain	
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	
If YES, with whom? <i>Rick McGowan, Benjamin M Scarborough, Jeroen Ruigrok van der Werven.</i>	
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	
Reference: <input checked="" type="checkbox"/> No	
4. The context of use for the proposed characters (type of use; common or rare)	
Reference: <input checked="" type="checkbox"/> Rare	
5. Are the proposed characters in current use by the user community?	
If YES, where? Reference: <i>They are found in some dictionaries and scholarly publications. See examples below.</i>	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	
If YES, is a rationale provided?	
If YES, reference: <i>Kana for WI, WE, and YE are all obsolete letters now. While WI and WE are already encoded in the BMP, YE is not yet. For further information, see below.</i>	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? <input checked="" type="checkbox"/> No	
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	
If YES, is a rationale for its inclusion provided?	
If YES, reference: <input checked="" type="checkbox"/>	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	
If YES, is a rationale for its inclusion provided?	
If YES, reference: <input checked="" type="checkbox"/> No	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?	
If YES, is a rationale for its inclusion provided?	
If YES, reference: <input checked="" type="checkbox"/>	
11. Does the proposal include use of combining characters and/or use of composite sequences? <input checked="" type="checkbox"/> No	
If YES, is a rationale for such use provided?	
If YES, reference: <input checked="" type="checkbox"/>	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference: <input checked="" type="checkbox"/>	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	
If YES, describe in detail (include attachment if necessary) <input checked="" type="checkbox"/> No	
<input checked="" type="checkbox"/>	
13. Does the proposal contain any Ideographic compatibility character(s)? <input checked="" type="checkbox"/> No	
If YES, is the equivalent corresponding unified ideographic character(s) identified?	
If YES, reference: <input checked="" type="checkbox"/>	

Summary

Hiragana and Katakana of the present-day do not include letters for "YE". This is because the "E" sound and the "YE" sound had merged in the 10th century in Japanese. In other words, "E" and "YE" were distinguished before, and were being represented by distinct Kana letters.

This is a proposal to add two missing Kana characters to the BMP of UCS for providing a way to distinguish in writing between "E" and "YE" mainly for Japanese classical literature.

Proposed Characters

1. The character "ゑ" (U+3048, HIRAGANA LETTER E) was derived from a cursive form of kanji "衣". This was being used to represent the "E" sound before "E" and "YE" merged. Therefore, the necessary Hiragana character is "HIRAGANA LETTER YE" in order to distinguish between "E" and "YE".



Shape of *HIRAGANA LETTER YE*

This character was derived from a cursive form of kanji "江", which was pronounced as "YE" when the distinction existed.

2. The character "ヰ" (U+30A8, KATAKANA LETTER E) was developed from parts (right half) of kanji "江". As mentioned above, this was being used to represent the "YE" sound originally. In the context that "E" and "YE" should be distinguished, **ヰ** should be used to represent its original sound "YE"; therefore the missing Katakana character is "KATAKANA LETTER ORIGINAL E".



Shape of *KATAKANA LETTER ORIGINAL E*

This character was developed from parts (first three strokes) of kanji "衣". As mentioned above, this character was pronounced "E".

E and YE (until the 10th century)

Pronunciation	e [e]	ye [je]
Original Kanji	衣	江
Hiragana	ゑ ^{*1}	ゑ ^{*3}
Katakana	ヰ ^{*4}	ヰ ^{*2}

*1 U+3048 HIRAGANA LETTER E

*2 U+30A8 KATAKANA LETTER E

*3 Proposed character: HIRAGANA LETTER YE

*4 Proposed character: KATAKANA LETTER ORIGINAL E

Proposed Code Points and Properties

3097;HIRAGANA LETTER YE;Lo;0;L;;;;N;;;;
????;KATAKANA LETTER ORIGINAL E;Lo;0;L;;;;N;;;;

Unfortunately, there seems already no empty space in the KATAKANA block, so that a new Katakana block, such as Katakana Extended-A, would be required for encoding KATAKANA LETTER ORIGINAL E. However, I personally would suggest that KATAKANA LETTER ORIGINAL E be encoded in U+3098 or U+3040 specially rather than in a new block created only for this one character, though both U+3098 and U+3040 are in the HIRAGANA block.

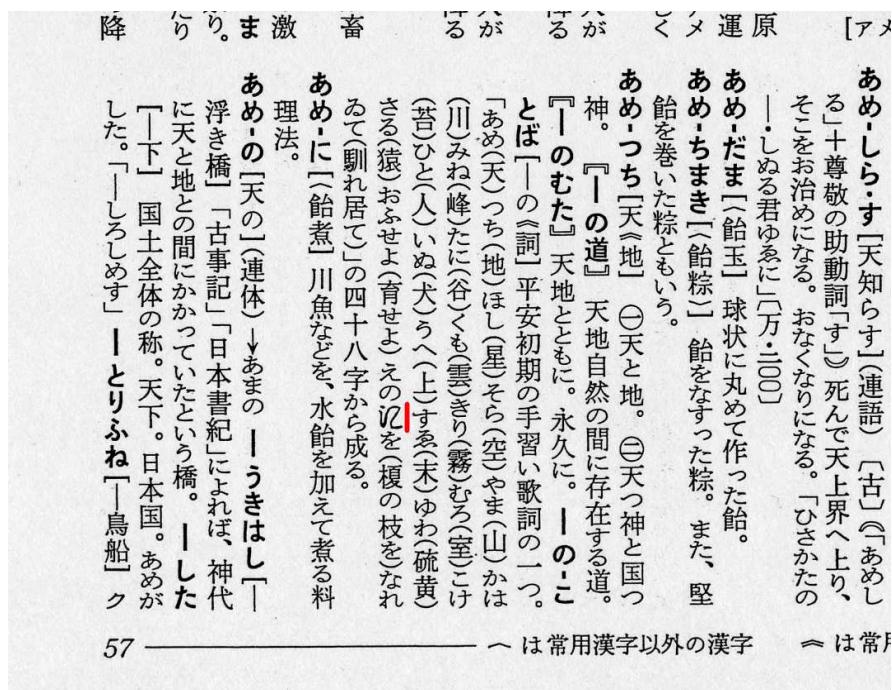
Sorting Order

1. HIRAGANA LETTER YE should be between U+3086 HIRAGANA LETTER YU and U+3087 HIRAGANA LETTER SMALL YO.
2. KATAKANA LETTER ORIGINAL E should be between U+30A6 KATAKANA LETTER U and U+30A7 KATAKANA LETTER SMALL E.

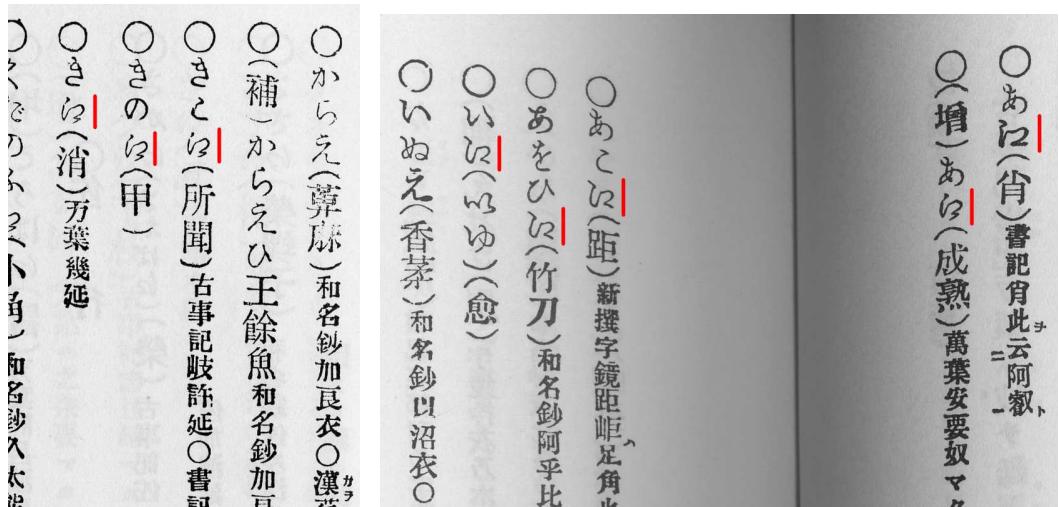
References

- Okumura Teruzane (奥村栄実), *Kogen E-Ye-ben* (古言衣延辨), the early 19th century, 1835, 1891.
 - Ooya Tooru (大矢透), *Kogen E-Ye-ben Shōho* (古言衣延辨證補), 1908.
- These were bound and issued by Benseisha (勉誠社) in 1977, now Bensey Publishing Inc. (勉誠出版).

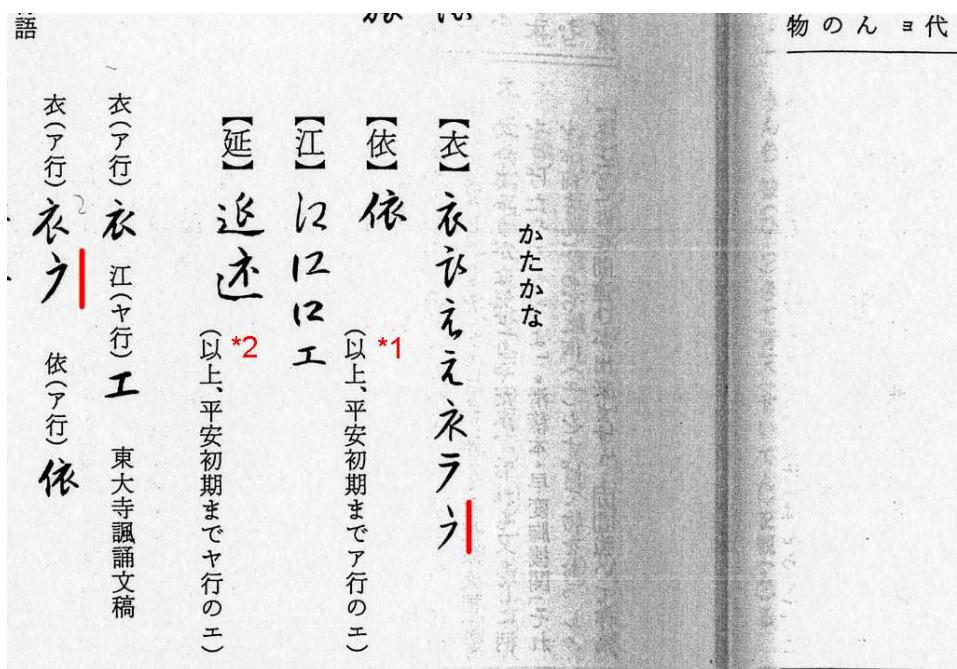
Examples



1. An example of HIRAGANA LETTER YE used in a dictionary, from *Kōjirin* (広辞林) the 6th edition, Sanseidō (三省堂), 1983, 1984, 1991.



2. Examples of HIRAGANA LETTER YE, from Okumura Teruzane (奥村栄実),
Kogen E-Ye-ben (古言衣延辨) printed version, 1891.



3. Development and variants of KATAKANA LETTER ORIGINAL E, from Nihon Kokugo Daijiten reduced-size edition (日本国語大辞典・縮刷版) volume 2, Shōgakukan (小学館), 1979, 1992.

*1: it is noted here that the right side characters were E until the early Heian period.

*2: it is noted here that the right side characters were YE until the early Heian period.

オ	オ	オ	オ	オ	オ	オ	オ	オ
オク〔置・在〕	オイヂウニ(懸重)	オイテ(據)	オイモチせず(隠没)	オ	オラブ(選)	オラビとりて(選取)	オラブ(選)	ナ(得)
奥14	二七12	一五17	二七12	一五23	ナ	ナラブ(選)	ナラビとりて(選取)	ナ(得)
オ	オ	オ	オ	オ	一五11	一五10	一五7	一五7

4. Examples of KATAKANA LETTER ORIGINAL E and HIRAGANA LETTER YE, from Kasuga Masaji (春日政治), *Saidaijibon Konkōmyō Saishōdōkyō Koten no Kokugogakuteki Kenkyū* (西大寺本金光明最勝王經古点の国語学的研究), Benseisha (勉誠社), 1985.

衆九万八千人ありキ。皆是レ阿羅漢なり。能ク善ク調伏せり。「イ調伏シ」
「イ如ク」諸の漏已に除し「イ除キ」、復煩惱無し。心善(く)解脱し「イシ」、惣
「イ解脱シ」所作已に畢へたり。「イ畢ヘ」諸の重擔を捨てて「イステ」、逮タリ
得タリ「イヲ」、諸の有の結を盡せり。「イ盡シ」、大自在を得たり。「イヲ得」、清淨
「イ住シ」善巧方便の智慧に莊嚴(セラレ)たり。「イラレ」、八解脱を證せり。
別句也
彼岸に到レリ。「イ到レルヲモチテ」其の名をば「日」具壽阿若橋陳如と、具
壽婆濕波と、具壽摩訶那摩と、具壽婆帝利迦と、大迦攝婆と、優樓頻螺迦
攝と、那提迦攝と、舍利子と、大目乾連といふ。唯し阿難陀のみ「於」學
是等の如キ諸の大聲聞あり。各「於」晡時に、定從(より)して「而」起(ち)て、佛
頂をモチテ佛の足を礼(したてまつり)て、右に三帯繞りて、退(き)て一面
復菩薩摩訶薩百の千万億の人有りて俱なり。大威德有(り)て、大龍王の
普ク聞(り)て衆に知し識セ所レたり。施と戒とをも清淨にすることを常に
し、忍行と精勤とをも無量の劫を経てし、諸に超(レ)たる靜慮に「イ諸の靜

Relevant Information

I have contacted Jeroen Ruigrok van der Werven, who has been preparing another proposal concerning Katakana. As of 14th December 2007, characters in each proposal do not cover nor conflict with the other. Our proposals could be combined, however, I would like to make it clear who is responsible for what characters, including fonts; so that I have chosen to submit my proposal to the Unicode Consortium separately.

<END OF PROPOSAL>