



DATE: 2007-04-20

ISO/IEC JTC 1/SC 2/WG 2

Universal Multiple-Octet Coded Character Set (UCS) - ISO/IEC 10646

Secretariat: ANSI

TITLE: Proposal to encode two Bopomofo characters in UCS
 SOURCE: TCA
 STATUS: Contribution from Liaison Member
 ACTION: with N3179, for consideration by WG2
 DISTRIBUTION: ISO/IEC JTC1/SC2/WG2

A. Administrative

1. Title: Proposal to encode two Bopomofo characters in UCS
2. Requester's name: TCA
3. Requester type (Member body/Liaison/Individual contribution): Liaison
4. Submission date: 2007/4/20
5. Requester's reference (if applicable): No
6. Choose one of the following:
 This is a complete proposal: Yes
 or, More information will be provided later: No

B. Technical - General

1. Choose one of the following:
 - a. This proposal is for a new script (set of characters): No
 Proposed name of script: _____
 - b. The proposal is for addition of character(s) to existing block: Yes
 Name of existing block: Bopomofo
2. Number of characters in proposal: 2
3. Proposed category (select one from below - see section 2.2 of P&P document):
 A-Contemporary X

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. Proposed Level of Implementation (1, 2 or 3) (see Annex K in P&P document): 1

Is a rationale provided for the choice? No

If Yes, reference: _____

5. Is a repertoire including character names provided? Yes

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

6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? the Chinese Foundation for Digitization Technology (CMEX)

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

7. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
 No

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

8. 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)? No

9. 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.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? 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.)? Yes

If YES, with whom? Bureau of Standard, Metrology and Inspection (BSMI)

If YES, available relevant documents: The contract between BSMI and CMEX

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Students in primary schools of Taiwan, Linguists

Reference: _____

4. The context of use for the proposed characters (type of use; common or rare): Bopomofo characters are daily used in all primary education of Taiwan and a lot of publications for children reading..

Reference: for example, <http://140.111.1.43/b3/body/3624.html>, an instance of the on-line dictionary on the MOE website, or <http://tw.dictionary.yahoo.com/search?ei=UTF-8&p=%E7%95%A2>

5. Are the proposed characters in current use by the user community? Yes

If YES, where? Reference: as mentioned above

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? Yes

If YES, is a rationale provided? Yes

If YES, reference: Keep with Bopomofo

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? No

8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? No

If YES, is a rationale for its inclusion provided? _____

If YES, reference: _____

9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? No

If YES, is a rationale for its inclusion provided? _____

If YES, reference: _____

10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? No

If YES, is a rationale for its inclusion provided? _____

If YES, reference: _____

11. Does the proposal include use of combining characters and/or use of composite sequences? No

If YES, is a rationale for its inclusion provided? _____

If YES, reference: _____

Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? No

If YES, reference: _____

12. Does the proposal contain characters with any special properties such as control function or similar semantics? No


If YES, describe in detail (include attachment if necessary): _____


13. Does the proposal contain any Ideographic compatibility character(s)? No

If YES, is the equivalent corresponding unified ideographic character(s) identified? ____







If YES, reference: _____

D. Proposal

There are two characters need to be proposed here for encoding. One is the character , as BOPOMOFO LETTER HI. And the other one is the

character , as BOPOMOFO LETTER YI. Regarding the requirement of the Bopomofo letter HI and evidence, since it has been explained in document N3179. It won't be mentioned here again.

As we know the UCS contains a number of Bopomofo characters in the Bopomofo block (U+3100-U+312F) and the Bopomofo Extended block (U+31A0-U+31BF). Therefore, this document is proposed here for encoding at U+312E, as BOPOMOFO LETTER YI.

Under the Microsoft Windows OS, it is queer that Character  and Character  are using the same code, U+3127. But in the UCS, the code, U+3127, is mapping to Character . And there is no other code mapping to the Character . In Taiwan it is more easy and common to see and to use the Character  than the Character , especially in the field of education or publications

Evidence for this character like in the Chinese dictionaries which are call “國民常用標準字典”(ISBN957-09-0118-7) and “正中形音義綜合大字典”(ISBN 957-09-1623-0), both of these two dictionaries embodied many Chinese words which need the characters to be part of the phonogram. See figure 1 and 2.

We also can easily to see this character be printed on much kind of publications. For examples, the publication of the Ministry of Education in Taiwan, or the Mandarin daily news(國語日報).
See figure 3, 4 and 5.

銑	鯨	擻	擻	擻	憲	夕	廐	懔	嶮	匙	丁	廡	鹹	鵬	臻	閑	閒	銜	軾	賢	讖	誠	滋
八三	六六	六三	五九	五七	五二	四八	四八	四六	四四	三七	丁	三三	三三	三三	三三	二九	二九	一九	一八	一七	一五	一五	一六
鱸	鮮	羈	顯	險	鑿	銑	跪	諗	蜺	鮮	銜	筮	筮	省	癖	獾	獾	獾	獾	獾	獾	獾	獾
三五	三五	三五	二六	二二	一九	一九	二七	二七	二五	二五	二五	二三	二三	二二	二〇	二〇	二〇	二〇	二〇	二〇	二〇	二〇	一六
名	腺	羨	縣	線	縱	硯	眼	獸	獻	瀆	涎	混	現	旋	攔	惆	憲	峴	峴	先	佃	倪	丁
一四	一九	一三	一三	一三	二九	二四	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	丁
忻	心	廢	姚	姚	給	給	鮮	鮮	鮮	鮮	限	鐵	鈞	軒	賢	謙	謙	謙	見	衮	規	覓	丁
四〇	四四	四四	三三	三三	三三	三三	二六	二六	二六	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	丁
白	信	丁	鐸	鐸	鐸	馨	駮	鏊	鏊	辛	訢	新	莘	苾	瘁	牲	焮	焮	歆	欣	忻	新	丁
二五	二二	丁	一五	一五	一五	二六	二五	一九	一九	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	丁
裏	蕻	菘	腳	壤	鄉	箱	相	皂	環	湘	復	廂	啞	丁	汎	蒙	苾	豐	罨	璽	瑗	熏	焮
一六	一五	一四	一七	一七	二二	二二	二二	一九	二二	二二	二二	二二	二二	二二	一九	一九	二二	二二	二二	二二	二二	二二	二二
蕻	珣	鼎	攘	想	嚮	亨	亨	丁	降	降	詳	虹	翔	祥	棒	棒	庠	丁	驤	香	鑲	鄉	丁
二五	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二
鄉	象	祿	術	蟻	蠆	相	珣	珣	椽	鼎	巷	嚮	向	像	丁	養	饗	饗	餉	響	鄉	壤	蠆
一八	一七	一五	一五	一五	一五	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二
嫻	型	刑	例	丁	鯉	驛	銜	鮮	蛭	興	腥	腥	理	狸	狸	星	惺	丁	鑲	項	關	關	鉅
三三	二五	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二
杏	悻	性	幸	倖	姓	倖	倖	丁	醒	筭	省	消	擗	丁	陘	鋼	鉞	邢	行	肝	斫	形	嶮
六九	五〇	四一	四〇	三六	三五	三五	三五	三五	二八	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二
吁	攬	戍	孀	嬖	妯	墟	嘔	嘔	煦	响	吁	暉	隹	丁	魁	諛	行	巷	苻	興	甕	勝	紆
六九	五九	五五	五三	五三	五三	二八	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二
魑	鬚	驢	須	需	鑄	誦	訃	襦	婿	虛	蕪	芋	胥	綸	綸	緝	稽	吁	弃	獻	歆	楛	跨
二五	二四	二四	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二	二二

注音符號檢字表

Figure 1. one of the table in the “國民常用標準字典”



Figure 3. Publication of the Ministry of Education in Taiwan (1)

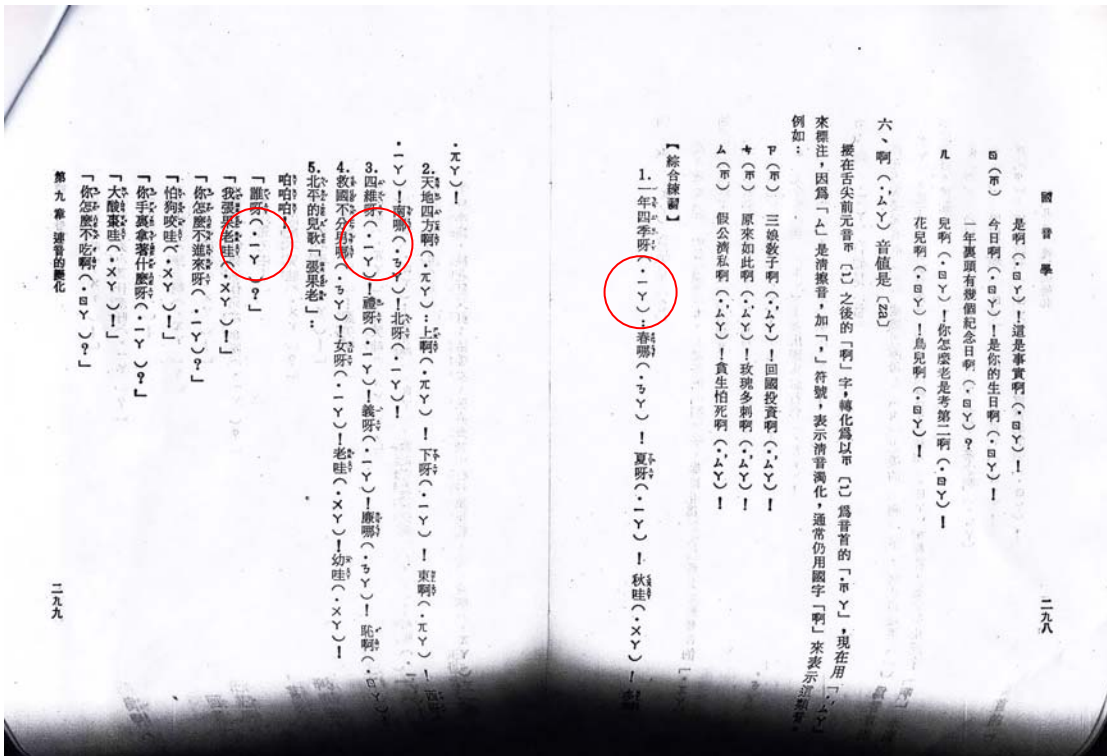


Figure 4. Publication of the Ministry of Education in Taiwan(2)



Figure 5. the Mandarin daily news(國語日報)

總表

ㄅ	ㄆ	ㄇ	ㄏ	ㄏ
ㄉ	ㄊ	ㄋ	ㄌ	
ㄍ	ㄎ	ㄎ	ㄏ	
ㄐ	ㄑ	ㄑ	ㄒ	
ㄓ	ㄔ	ㄔ	ㄕ	
ㄖ	ㄗ	ㄗ	ㄘ	ㄘ
ㄙ	ㄚ	ㄚ	ㄛ	
ㄜ	ㄝ	ㄝ	ㄞ	ㄞ
ㄟ	ㄠ	ㄠ	ㄡ	ㄡ
ㄣ	ㄤ	ㄤ	ㄥ	ㄥ
ㄦ	ㄨ	ㄨ	ㄩ	ㄩ
ㄩ	ㄩ	ㄩ	ㄩ	ㄩ
ㄩ	ㄩ	ㄩ	ㄩ	ㄩ

- 註：1. 聲符「万」、「兀」、「广」國音不用。
 2. 韻符「帀」注音時省略不標。
 3. 韻符「一」直式注音寫成「一」，橫式注音寫成「丨」。