

Unicode request for ɪ with retroflex hook

Kirk Miller, kirkmiller, gmail.com

2025 May 09

In [L2/20-125](#), a number of characters with retroflex hook were requested, including vowel letters where the hook was used to indicate assimilation to a retroflex consonant. Since then the letter ɪ has been attested for the retroflex *shi* vowel of Mandarin (Chu & Petrus 2020). Petrus (p.c. 2025) notes that a number of dissertations have referenced this transcription. Luo (2024: 18), for example, states that they've adopted the Chu & Petrus transcription. However, they substitute <ɪ> for <ɪ̌>, apparently due to a lack of font support; Petrus (p.c. 2025) explains that [ɪ̌] is not the same sound as [ɪ] and so is not an adequate substitute, especially for singing.

ɪ̌ ɪ̌	ɪ̌ ɪ̌	ɪ̌ ɪ̌
ɪ̌ ɪ̌		ɪ̌
ɪ̌ ɪ̌	ɪ̌ ɪ̌	ɪ̌ ɪ̌
ɪ̌ ɪ̌	ɪ̌	ɪ̌ ɪ̌
ɪ̌	ɪ̌	ɪ̌ ɪ̌
	ɪ̌	

Table 1. Retroflex-hook modifications of the vowel letters on the 1979 IPA chart, with the alternative variants of the vowels ɪ̌ and ɪ̌. Prior to this time the retroflex hook was a way to write rhotic vowels; now it is used to indicate assimilation. Letters in black are encoded; the requested letter is highlighted in red.

Character

Letter with retroflex hook

ɪ̌ 1DF85 LATIN LETTER SMALL CAPITAL I WITH RETROFLEX HOOK.

Properties

1DF85;LATIN LETTER SMALL CAPITAL I WITH RETROFLEX HOOK;L;0;L;;;;N;;;;;

DoNotEmit data

026A 0322; 1DF85; Precomposed_Form # LATIN LETTER SMALL CAPITAL I, COMBINING RETROFLEX HOOK BELOW; LATIN LETTER SMALL CAPITAL I WITH RETROFLEX HOOK

References

Katherine Chu & Juliet Petrus (2020) *Singing in Mandarin: A Guide to Chinese Lyric Diction and Vocal Repertoire*. Rowman & Littlefield.

Chuyan Luo (2024) *A Study of Selected Vocal Works by Shande Ding: The Mysterious Sound of Flute and Poems on Western Yunnan*. PhD dissertation, Louisiana State University.

Chart

Greyed out cells are assigned (medium grey) or accepted by the UTC or SEW (light grey).

1DF00

Latin Extended-G

1DFFF

	1DF0	1DF1	1DF2	1DF3	1DF4	1DF5	1DF6	1DF7	1DF8	1DF9	1DFA	1DFB	1DFC	1DFD	1DFE	1DFF
0	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
1	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
2	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
3	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
4	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
5	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
6	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
7	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
8	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
9	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
A	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
B	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
C	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ					ṡ	ṡ	ṡ
D	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ				ṡ	ṡ	ṡ	ṡ
E	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ				ṡ	ṡ	ṡ	ṡ
F	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ	ṡ				ṡ	ṡ	ṡ	ṡ

Figures

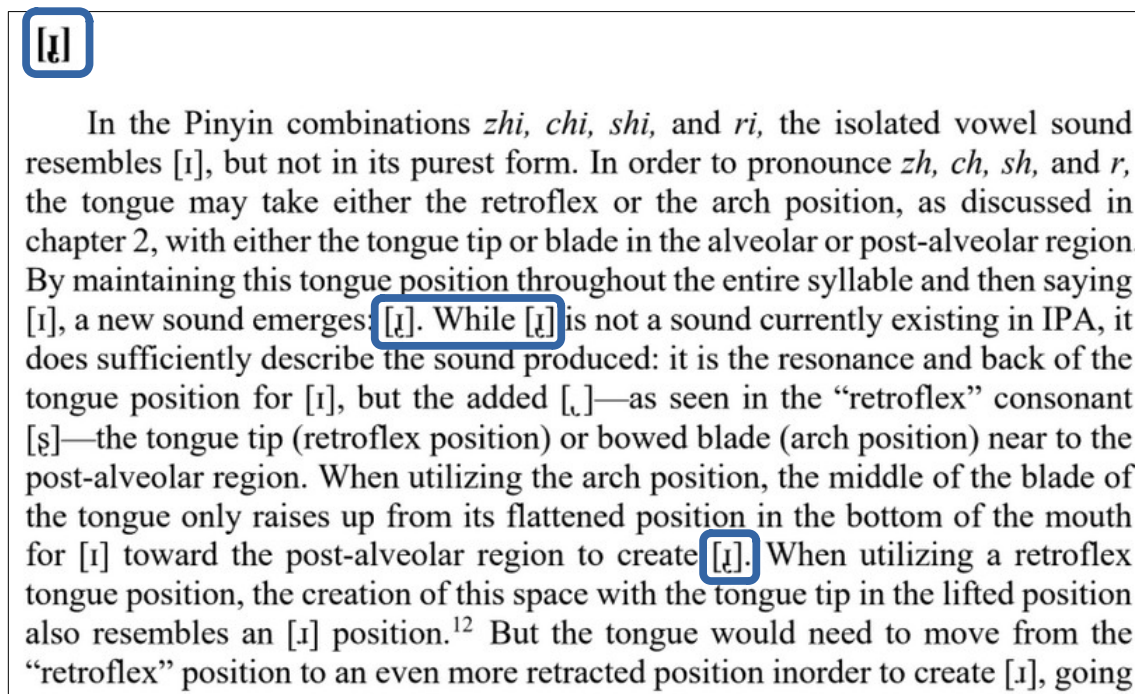


Figure 1. Chu & Petrus (2020: 46). A description of <ɿ̚>. It is used for the retroflex syllables written *zhi*, *chi*, *shi*, *ri* in pinyin, whereas U+1D7B <ɿ> LATIN SMALL CAPITAL LETTER I WITH STROKE is used for non-retroflex *zi*, *ci*, *si*. The retroflex hook is used to indicate that the vowel is pronounced with the same tongue position as the preceding consonant.

🎧 Exercise 3.2C: Practice speaking and then singing the following syllables containing [ʂ̚], [tʂ̚], [tʂ̚ʰ], and [ʉ̚]. Try them both as short and long sounds to practice the sustaining of the vowel. Remember to not change the mouth and tongue position when moving from the consonant to the apical vowel. Say, then stay.

只 <i>zhǐ</i>	[tʂ̚]	(only)	知 <i>zhī</i>	[tʂ̚]	(to know)
直 <i>zhí</i>	[tʂ̚]	(direct, straight)	吃 <i>chī</i>	[tʂ̚ʰ]	(to eat)
持 <i>chí</i>	[tʂ̚ʰ]	(to grasp)	迟 <i>chí</i>	[tʂ̚ʰ]	(late; tardy)
是 <i>shì</i>	[ʂ̚]	(to be)	时 <i>shí</i>	[ʂ̚]	(hour)
诗 <i>shī</i>	[ʂ̚]	(poem)	日 <i>rì</i>	[ʉ̚]	(day; sun)

Figure 2. Chu & Petrus (2020: 47). <ɿ̚> in transcription. The placement of the hook is a typographic choice. It presumably appears on the left serif here due to poor font support of the combining diacritic U+0322 ◌̚ on a narrow letter, as seen in Noto Serif <ɿ̚>. That may look good in <ʂ̚> as seen here but would clash in <ʂ̚̚>. Default placement in IPA is on the right, as in <ɿ̚> and <ɿ̚̚>.

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 std.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 std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html.
See also std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest Roadmaps.**

A. Administrative

1. Title:	<i>I with retroflex hook</i>
2. Requester's name:	<i>Kirk Miller</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>individual</i>
4. Submission date:	<i>2025 May 09</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>yes</i>
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	<i>no</i>
Proposed name of script:	
b. The proposal is for addition of character(s) to an existing block:	<i>yes</i>
Name of the existing block:	<i>Latin Extended-G</i>
2. Number of characters in proposal:	<i>1</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>
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?	<i>yes</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>yes</i>
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Kirk Miller</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>SIL (Gentium release)</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>yes</i>
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)?	
	<i>no</i>

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 Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N4502-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, 2008-05, 2009-11, 2011-03, 2012-01)

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.)?	no
If YES, with whom?	
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:	
4. The context of use for the proposed characters (type of use; common or rare)	phonetic
Reference:	
5. Are the proposed characters in current use by the user community?	yes
If YES, where? Reference: See references	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	no
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	NA
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, or could be confused with, an existing character?	yes
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 such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
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 characters?	no
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	