

Unicode request for zh ligature

Kirk Miller, kirkmiller, gmail.com

2026 February 06

This is an addendum to [L2/24-273](#) *Unicode request for Initial Teaching Alphabet* for an additional letter to the alphabet: the zh ligature <zh>.

In recent literacy projects, teachers objected that some of the ITA letters were not close enough to the standard alphabet, so the glyphs were adjusted in accommodation. (Compare Figure 4 with Figure 5.) In some cases this merely enshrined allographs already present in the 1960s, such as <æ> and <œ> with a short linking stroke instead of the common ASCII substitutes <æ> and <œ>, as illustrated in L2/24-273. Others were new designs, such as <ch> and <ee> for older <h> and <e>. These can be treated as allographs of the original letters and be left to the font. However, one letter cannot be treated this way: The /z/ sound had originally been written <3> as in IPA, and was identified with the existing Unicode character U+0292 3, but it has now been replaced with the unsupported ligature <zh>. We therefore request this new character. Doing so does not interfere with the adoption of L2/24-273, as the original ITA letter <3> had not been requested in that proposal.

Characters

zh 1DF98 LATIN SMALL LETTER ZH.

Properties

1DF98;LATIN SMALL LETTER ZH;Ll;0;L;;;;N;;;;

Annotations

1DF98 LATIN SMALL LETTER ZH

- in ITA replaces older letter U+0292 3

Collation

ITA material is typically alphabetized in standard orthography and then converted to ITA. It would make sense therefore to collate the characters as digraphs rather than as unrelated symbols. That is, requested U+1DF98 zh should be treated as zh, and similarly the U18 characters ch ee ie oi ou ω sh th th wh (allographically ch ee ie oi ou ω sh th th wh) as ch ee ie oi ou oo sh th (× 2) and wh digraphs.

References

- Jane Anderson (2017) *ITA Linguistic Readers, book 1*. Initial Teaching Alphabet Foundation, New York.
- (2025) *ITA First Readers, book 1: The numbers kids*. Initial Teaching Alphabet Foundation, New York.
- Thomas Spence (1775) *The Grand Repository of the English Language*. Newcastle upon Tyne.

Charts

ITA characters are highlighted, including those in the Pipeline for Unicode 18.

Latin Extended-G

1DF00

1DFFF

	1DF0	1DF1	1DF2	1DF3	1DF4	1DF5	1DF6	1DF7	1DF8	1DF9	1DFA	1DFB	1DFC	1DFD	1DFE	1DFF
0	ᄁ	ᄂ	ᄃ	ᄄ	ᄅ	ᄆ	ᄇ	ᄈ	ᄉ	ᄊ				ᄋ	ᄌ	ᄍ
1	ᄎ	ᄏ	ᄐ	ᄑ	ᄒ	ᄓ	ᄔ	ᄕ	ᄖ	ᄗ				ᄘ	ᄙ	ᄚ
2	ᄛ	ᄜ	ᄝ	ᄞ	ᄟ	ᄠ	ᄡ	ᄢ	ᄣ	ᄤ				ᄥ	ᄦ	ᄧ
3	ᄨ	ᄩ	ᄪ	ᄫ	ᄬ	ᄭ	ᄮ	ᄯ	ᄰ	ᄱ				ᄲ	ᄳ	ᄴ
4	ᄶ	ᄷ	ᄸ	ᄹ	ᄺ	ᄻ	ᄼ	ᄽ	ᄾ	ᄿ				ᄿ	ᄿ	ᄿ
5	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
6	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
7	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
8	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
9	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
A	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
B	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
C	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
D	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
E	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ
F	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ	ᄿ			ᄿ	ᄿ	ᄿ	ᄿ

Figures

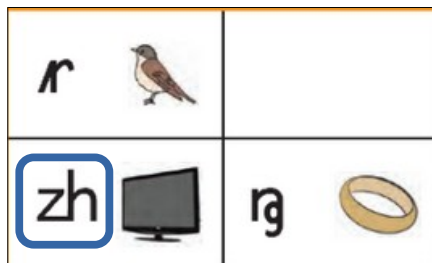


Figure 1. Anderson (2025: 111). <zh> in an alphabet chart.

then wee wach teluvizhun.
lilee and anu wach teluvizhun.
ie wach teluvizhun, tω.

Figure 2. Anderson (2017: 28). <zh> in text.

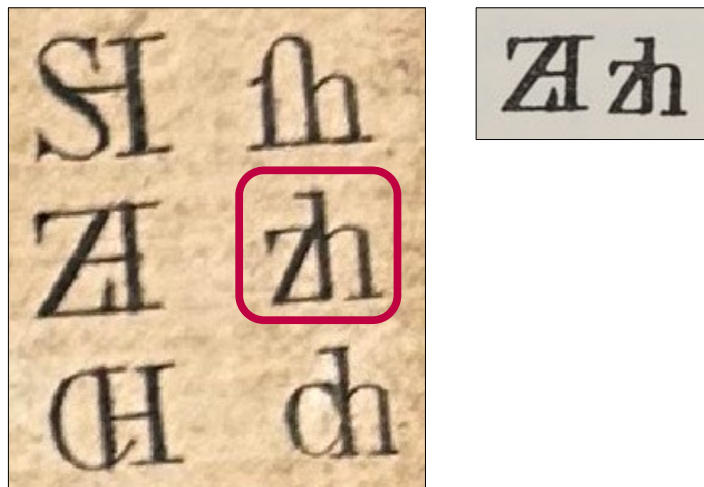


























Figure 3. Spence (1775). The expected generic form of the ligature. Although the glyph in Anderson is open at the top: <zh>, the same is true of ʒh: <ch>, and this is not reflected in the Unicode chart. I therefore propose a simple ligature <zh>, as seen here, as the representative glyph.

a 	æ 	b 	h 	n 	v 
e 	ee 	c 	j 	p 	w 
i 	ie 	d 	k 	r 	y 
o 	oe 	f 	l 	s 	z 
u 	ue 	g 	m 	t 	









































ɑ 	au 	r 	z 	sh 	th 
ω 	ω 	3 	rg 	ch 	th 
ou 	oi 			wh 	

Figure 4. The typeface used from 1993 to present by the ITA Foundation of New York, with <3>, and the alphabet chart currently posted on their website at *What is ITA?*. This was a source for the ITA glyphs currently in the pipeline for Unicode 18.

a 	æ 	b 	h 	n 	v 
e 	ee 	c 	j 	p 	w 
i 	ie 	d 	k 	r 	y 
o 	oe 	f 	l 	s 	z 
u 	ue 	g 	m 	t 	















ɑ 	au 	r 		sh 	th 
ω 	ω 	zh 	rg 	ch 	th 
ou 	oi 			wh 	

Figure 5. Anderson (2017: 54, 2025: 111) A second typeface used by the ITA Foundation since at least 2017 for a literacy project in Puerto Rico. <a> and <ɑ> are swapped, and the ligatures look more like standard Latin letters. <3> is replaced by the requested ligature <zh>.

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>Zh ligature</i>
2. Requester's name:	<i>Kirk Miller</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>individual</i>
4. Submission date:	<i>2026 February 06</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"/>	D-Attested 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 www.unicode.org for such information on other scripts. Also see Unicode Character Database (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.)?	yes
If YES, with whom? ITA Foundation	
If YES, available relevant documents: emails available	
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)	phonics
Reference:	
5. Are the proposed characters in current use by the user community?	yes
If YES, where? Reference:	
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)?	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, or could be confused with, 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 such use 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 characters?	no
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	