

Glyph corrections for U+027F and U+0285 in TUS

Lorna A Priest and Greg Aumann
SIL International
 2007-Jan-12

Introduction

The charts for The Unicode Standard show U+027F LATIN SMALL LETTER REVERSED R WITH FISHHOOK and U+0285 LATIN SMALL LETTER SQUAT REVERSED ESH with the glyphs on the baseline (see Figure 1). Chinese books consistently use these characters with the glyph descending **below** the baseline as shown below and in Figures 2-7. Figures 2b and 4b are a magnification of one portion of Figures 2a and 4a respectively to show clearly where the baseline is in relation to these characters. We would like to request the UTC to correct the glyphs in the charts and elsewhere where they may be referenced.

baseline ʀ ʀ baseline

The symbols were invented by Karlgren about 100 years ago. Figure 2b shows how he used the characters with the glyph descending below the baseline.

Both characters were added to Unicode 1.1. The shape of U+0285 LATIN SMALL LETTER SQUAT REVERSED ESH was changed between version 2.0 and 3.0 of the Unicode standard to be “squat.” Since the name has not changed it’s probable the original samples for U+0285 were “squat” and the glyph in TUS was just wrong from the beginning.

We have been unable to find the document requesting the glyph change for U+0285 or the original document proposing these characters. If the original proposal used older printed materials the baselines do jump around and if there are no characters immediately adjacent it would be very difficult to tell where the baseline is intended to be. The examples in this document are words with lots of adjacent characters and recent versions so it is much clearer where the baseline is.

Figures

5	ċ	ȳ	ø	ɹ	ʃ	ɟʒ
	0255	0265	0275	0285	0295	02A5
6	đ	ħ	œ	ƒ	ʝ	ʦ
	0256	0266	0276	0286	0296	02A6
7	ɖ	ɧ	ω	ɽ	ç	ʧ
	0257	0267	0277	0287	0297	02A7
8	ə	ɨ	ϕ	ɾ	⊙	ɽ
	0258	0268	0278	0288	0298	02A8
9	ə	ɹ	ɹ	ɥ	ʙ	ɧ
	0259	0269	0279	0289	0299	02A9
A	æ	ɪ	ɪ	ʊ	ɛ	ɬ
	025A	026A	027A	028A	029A	02AA
B	ɛ	ɨ	ɹ	ʊ	ɣ	ɬ
	025B	026B	027B	028B	029B	02AB
C	ɜ	ɨ	ɹ	ʌ	ɥ	ɰ
	025C	026C	027C	028C	029C	02AC
D	ɜ	ɹ	ɾ	ɹ	ɨ	ɰ
	025D	026D	027D	028D	029D	02AD
E	ɛ	ɧ	ɹ	ɹ	ɣ	ɹ
	025E	026E	027E	028E	029E	02AE
F	ɨ	ɹ	ɹ	ɹ	ɹ	ɹ
	025F	026F	027F	028F	029F	02AF

Figure 1. TUS. p. 591.

ɪ ɪ, voyelle apico-gingivale, haute, tendue, délabialisée ou à l'ouverture labiale large, se rencontre en Mand. Yangtc. Wou Yue, peut-être Swat.¹ Jap. Elle est partout orale, et n'apparaît qu'en syllabe ouverte et après s, z, p. ex. Pék. sɪ, hch. sseu 'mourir'. Edkins compare cette voyelle au 'e' angl. de tassel, comparaison assez heureuse en considérant que dans ce mot la langue, à cause du ɪ suivant, ne s'éloigne que très peu de la position du s. L'absence de ɪ dans les autres langues connues rend sa définition très difficile pour des amateurs. Les identifications courantes comme « l'u bref anglais », « l'eu français » etc. sont toutes très incorrectes.

ɨ, voyelle apico-alvéolaire, haute, tendue, délabialisée ou à l'ouverture labiale large existe en Pék. Chansi (sauf T'aiy. T'aik. et P'ingy.) Kansou Chensi Hoaik. K'aif. Sseu. Nank. Elle est partout orale et n'existe qu'en syllabe ouverte après s, z, p. ex. Pék. sɨ, hch. che 'temps'. C'est probablement aussi cette voyelle qu'a eu en vue M. Parker, quand il écrivait chī, chīng pour le dial. de Ningpo. Ce qui est dit ci-dessus des définitions de la voyelle ɪ s'applique aussi bien à celles de ɨ, publiées jusqu'ici. Une voyelle très analogue à ɨ se rencontre dans certains dial. suédois; on l'appelle « i de Viby ».

Les voyelles ɪ et ɨ ont causé de l'embarras aux transpositeurs. Les russes ont choisi la lettre ɪ pour toutes les deux, graphie assez bonne au point de vue acoustique. Le parti pris par M. Mateer, qui écrit 'i' pour toutes les deux est également bon. La

¹ Cf. ɪ de la série vélo-palatale (p. 308).

Figure 2a. Karlgren. p. 295.

Les voyelles ɪ et ɨ ont causé de l'embarras aux transpositeurs. Les russes ont choisi la lettre ɪ pour toutes les deux, graphie assez bonne au point de vue acoustique. Le parti pris par M. Mateer, qui écrit 'i' pour toutes les deux est également bon. La

Figure 2b. Karlgren. p. 295 (showing baseline).

Table 2. Finals

2(a) In G. R.

Ending Column	Open Ending			-i		-u		-n		-ng			-l
	1	2	3	4	5	6	7	8	9	10	11	12	13
Medial													
Row-a	y	a	e	ai	ei	au	ou	an	en	ang	eng	ong	el
Row-i	i	ia	ie	iai		iau	iou	ian	in	iang	ing	iong	
Row-u	u	ua	uo	uai	uei			uan	uen	uang	ueng		
Row-iu	iu		iue					iuan	iun				

2(b) In IPA

Ending Column	Open Ending			-i		-u		-n		-ŋ			-l
	1	2	3	4	5	6	7	8	9	10	11	12	13
Medial													
Row-a	ɿ, ʅ [*]	ɿ̚	ɿ̚	ai	ei	au	ou	an	ən	aŋ	ɤŋ	oŋ	ɿ̚
Row-i	i	i̥	i̥	iai		iau	iou	iɛn	in	iaŋ	iŋ	iɔŋ	
Row-u	u	u̥	u̥	uai	uei			uan	uən	uaŋ	uɤŋ		
Row-iu	y		yɛ					yan	yn				

*Karlǵren invented the symbols ɿ and ɿ̚ for these apical vowels, as he calls them. See his *Phonologie*, pp. 295-297, Chinese transl., pp. 197-199.

Figure 3. Chao. 1968. p. 24.

29. tʂ'	北京:痴[tʂ'ɿ]	43. ø	北京:希[øi]
30. dz	湖南双峰:池[dzɿ]	44. z	浙江绍兴:自[zɿ]
31. ʂ	北京:诗[ʂɿ]	45. k	北京:该[kai]
32. z	北京:日[zɿ]	46. k'	北京:开[k'ai]
33. c	浙江永康:居[cy]	47. g	苏州:茄[gɔ]
34. c'	浙江永康:区[c'y]	48. ŋ	苏州:误[ŋəu]
35. ɟ	浙江永康:渠[ɟy]	49. x	北京:海[xai]
36. ɲ	浙江永康:元[ɲyǝ]	50. ɣ	湖南双峰:鞋[ɣa]
37. ɕ	浙江永康:虚[ɕy]	51. x	山西太谷:胡[xu]
38. j	湖南韶山:夏[jia]	52. ʔ	云南玉溪:肝[ʔǝ]
j	广州:夜[je]	53. ʔ'	云南玉溪:宽[ʔ'uǝ]
39. tɕ	北京:鸡[tɕi]	54. h	苏州:花[ho]
40. tɕ'	北京:欺[tɕ'i]	55. ɦ	苏州:鞋[ɦɔ]
41. dz	苏州:骑[dzi]	56. 清音化: b, d.	
42. ɲ	苏州:银[ɲin]		

四、元音举例

1. ɿ	北京:丝[sɿ]	18. u	湖南双峰:波[pɔ]
2. ɿ	北京:诗[ʂɿ]	19. u	北京:乌[u]
3. y	苏州:书[sy]	20. y	北京:雨[y]
4. y	湖北麻城:树[sy]	21. ø	苏州:干[kø]
5. i	北京:衣[i]	22. œ	广州:靴[hœ]
6. i	苏州:烟[i]	23. ɤ	北京:饿[ɤ]
7. e	北京:梅[mei]	24. u	兰州:二[yu]
8. ɛ	苏州:三[sɛ]	25. ɛ	北京:儿[ɛ]
9. ɛ	广州:爹[tɛ]	26. ø	广州:虚[hø]
10. œ	苏州:好[hœ]	27. ~	[鼻化, 厦门:影[ĩǝ]
11. a	北京:干[kan], 广州:花[fa]	28. o ₁	舌较高
12. a	北京:家[tɕia]	29. o ₂	舌较低
13. ɐ	广州:闭[pɐ]	30. o ₃	唇较圆
14. a	北京:纲[kɑŋ]	31. o ₄	唇较展
15. ɔ	苏州:街[kɔ], 南京:茶[tɕɔ]	32. ɔ	开口度较小
16. ɔ	广州:火[fɔ]	33. o ₅	舌位较前
17. o	苏州:茶[zɔ]	34. o ₆	舌位较后

Figure 4a. Yuan, 1989. p. 6.

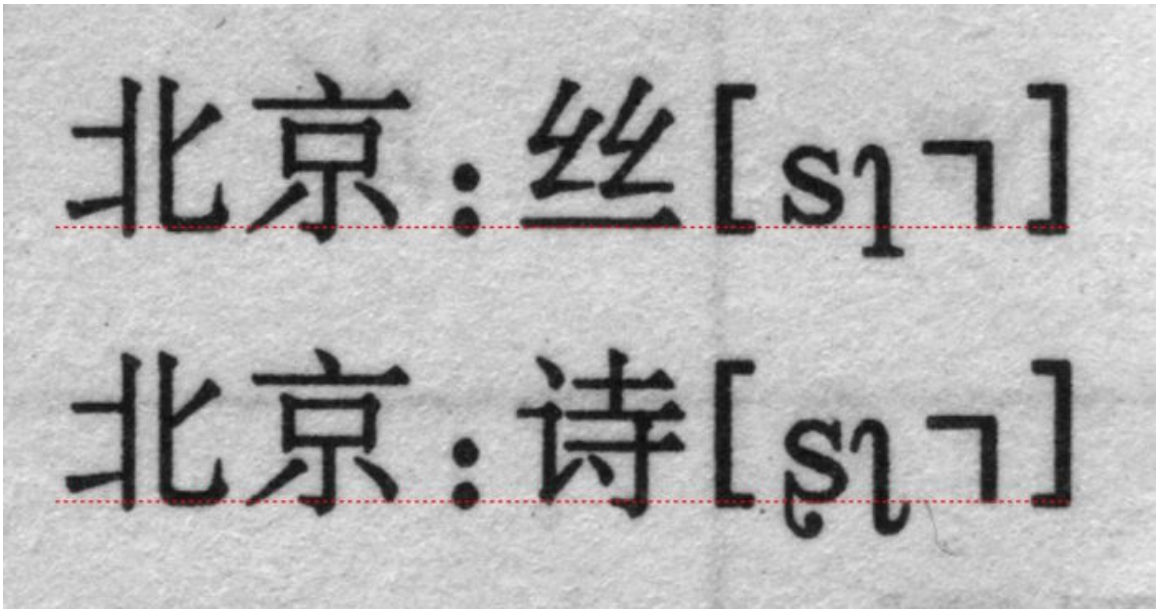


Figure 4b. Yuan. 1989. p. 6 (magnification showing baseline).

8. 两种来源的 pf-, pf'-

pf-, pf'-声母出现在西北。按来源可以分两种情况。

(1) 帮组来源的 pf-, pf'- ——陕西宝鸡、麟游等地, u 韵和 o 韵之前没有 p-, p'-, 只有 pf-, pf'-。如“跛”pfo, “坡”pf'o, “布”pfu, “朴”pf'u。^①

(2) 知照组来源的 pf-, pf'- ——凡北京 ts-, ts'-声母的 u 韵字, 在兰州、西安、潼关等地都是 pf-, pf'-声母的 u 韵字; 凡北京 ts-, ts'-声母的 u 韵以外的合口字, 在上述地区都是 pf-, pf'-声母的开口字(介音 u 失落), 如西安“猪”pfu, “初”pf'u; “追”pfei, “吹”pfei。不但如此, 凡北京的 su, zu, su-, zu-, 在上述地区都相应变为 fu, vu, f-, v-, 如西安“书”fu, “如”vu, “水”fei, “软”væ。但知照组来源的开口字, 在 ɿ 和 ʝ 二韵前读 ts- ts'-, 在其他韵母前读 t- t'-, 例如“知”tsɿ, “遮”tsʝ, “吃”tsɿ, “车”tsʝ, “朝”tau, “张”taŋ, “超”t'au, “昌”t'aŋ。但读书音中 t- t'-也已变成 ts- ts'-, 二者的区别渐趋泯灭。^②

山西永济、运城也有 pf-, pf'-声母,^③ 情况与西安相似。

9. “资雌私”和“知蚩诗”声母的分混

这两组字的声母在北京话是有分别的, “资雌私”是 ts-, ts'-, s-, “知蚩诗”是 tʃ-, tʃ'-, ʃ-。多数地区这两组声母有分别, 只是分别的具体情况, 即哪些字属 ts-组, 哪些字属 tʃ-组, 不全一样。完全不分的地区比较少些。大体上说, 北方方言中, 除东北以不分占优势(如沈阳、丹东、哈尔滨)外, 一般是能分的。西北方言中, 太原、沁县等地不分, 其他很多地区能分。西南方言绝大部分地区不分, 少数地区如昆明、楚雄能分。江淮方言一般不分, 南京、巢县等地能分。

各种分混情形列表如下(以 s-和 ʃ-的分混为例):

	送(心母)	生(审二)	诗(审三)	身(审三)
北京	suŋ ³	ʃəŋ	tʃɿ	ʃəŋ
当阳(湖北)	suŋ ³	ʃəŋ	tʃɿ	ʃəŋ
昆明	soŋ ³	sə	tʃɿ	ʃə
天津	suŋ ³	səŋ	tʃɿ	ʃəŋ
成都	soŋ ³	səŋ	tʃɿ	səŋ

10. 日母在各地的变化

日母字声母(这里不包括止撮开口三等字“儿耳二”等的读音)大部分地区一律是 z-, 或者一律是 ʒ-。一般是有 ts-组的地区是 z-, 如北京, 否则是 ʒ-。

有些地区情况不同。比如湖北当阳为 ʒ-; 扬州“热”是 ʒ-, 其他一律是 l-; 沈阳为 ʒ- (齐齿或撮口); 济南开口韵前为 z-, 合口韵前为 l-; 汉口合口撮口韵前和“日”是 ʒ-, 开口韵前是 n-; 西安、兰州等地合口韵前是 v- (参看第 8 节), 开口韵前是 z- (但西安“日”ʒ-)。如下表:

	人	日	热	如	软
北京	zən	ʒɿ	zʝ	zɿ	zuan
成都	zən	ʒɿ	ze	zu	zuan

① 参见白涤洲、喻世长《关中方言调查报告》第 6 页, 中国科学院出版, 1954 年。
 ② 同上。
 ③ 根据高本汉《中国音韵学研究》第 174 页, 商务印书馆, 1940 年。

Figure 5. Yuan. 1989. p. 31.

方音点	字目 中古音	肆 ^{放肆}	似	祀	寺	伺	知 ^② 知道	蚩	支
		止开三 去至心	止开三 上止邪	止开三 上止邪	止开三 去志邪	止开三 去志邪	止开三 平支知	止开三 平支知	止开三 平支章
北 京		sɿ'	sɿ' ^① ʃɿ'	sɿ'	sɿ'	sɿ'	.tʃɿ	.tʃɿ	.tʃɿ
济 南		sɿ'	sɿ' ^① ʃɿ'	sɿ'	sɿ'	sɿ'	.tʃɿ	.tʃɿ	.tʃɿ
西 安		sɿ'	sɿ' ^① ʃɿ'	sɿ'	sɿ'	sɿ' sɿ'	.tʃɿ	.tʃɿ	.tʃɿ
太 原		sɿ'	sɿ'	sɿ'	sɿ'	sɿ'	.tʃɿ	.tʃɿ	.tʃɿ
武 汉		sɿ'	sɿ'	sɿ'	sɿ'	sɿ' tsɿ' _旧	.tʃɿ	.tʃɿ	.tʃɿ
成 都		sɿ'	sɿ'	sɿ'	sɿ'	sɿ'	.tʃɿ	.tʃɿ tʃɿ _俗	.tʃɿ
合 肥		sɿ'	sɿ'	sɿ'	sɿ'	sɿ'	.tʃɿ	.tʃɿ tʃɿ _俗	.tʃɿ
扬 州		sɿ'	sɿ'	sɿ'	sɿ'	sɿ'	.tʃɿ	.tʃɿ	.tʃɿ
苏 州		sɿ'	zɿ'	zɿ'	zɿ'	zɿ'	.tʃɿ	.tʃɿ	.tʃɿ
温 州		sɿ'	'zɿ	'zɿ	zɿ'	zɿ'	.tʃɿ	.tʃɿ	.tʃɿ
长 沙		sɿ'	sɿ' _文 sɿ' _白	sɿ' _文 tsɿ' _白	tsɿ'	sɿ'	.tʃɿ	.tʃɿ tsɿ'	.tʃɿ
双 峰		sɿ'	dzɿ'	dzɿ'	dzɿ'	dzɿ'	.tʃɿ	.tʃɿ	.tʃɿ
南 昌		sɿ'	sɿ'	sɿ'	sɿ'	sɿ'	.tʃɿ	.tʃɿ	.tʃɿ
梅 县		si'	sɿ'	sɿ'	ts'ɿ'	sɿ'	.tʃɿ _文 ti _白	.tʃɿ	.tʃɿ
广 州		ʃi'	'tʃi	tʃi'	tʃi'	tʃi'	.tʃi	.tʃi	.tʃi
阳 江		ɬei'	tʃei' _文 tʃei' _白	tʃei'	tʃei'	tʃei'	.tʃi	.tʃi	.tʃi
厦 门		su'	su' _文 sai' _白	su' _文 sai' _白	si'	su' _文 ts'i' _白	.ti _文 ^③ ti' _文	.ti	.tsi _文 ki _白
潮 州		sɿ'	'sɿ	'sɿ	zi'	sɿ' _文 ts'i' _白	ti' _文 tsai _白	.ti	.tsɿ
福 州		søy'	søy'	søy' _文 sai' _白	sei'	søy' _文 ts'ei' _白	.ti	.t'i	.tsie
建 瓯		si'	su'	su'	tsi'	si'	.ti	.ti	.tsi

①sɿ'为“是”训读。 ②又通“智”，*知义切，止开三去真知。 ③tsai_白。

Figure 6. Linguistics Section. 1989. p. 59.

Table 8.5. Development of the Middle Chinese final consonants in Mandarin dialects

	Middle Chinese	Peking	Yángzhōu	Xiān	Chéngdū	Lingbǎo
'three'	sâm	san ¹	sã ¹	sã ¹	san ¹	sã ¹
'umbrella'	sân:	san ³	sã ³	sã ³	san ³	sã ³
'mulberry'	sâng	saŋ ¹	saŋ ¹	saŋ ¹	saŋ ¹	saŋ ¹
'twelfth month'	lâp	la ⁵	naɿ ⁷	la ¹	na ²	la ¹
'pungent'	lât	la ⁵	naɿ ⁷	la ¹	na ²	la ¹
'drop'	lâk	luo ⁵	naɿ ⁷	luo ¹	no ²	luo ¹

Leaving aside the medials and those finals which contain a single high vowel, most Mandarin dialects show a basic vocalic contrast between two degrees of openness, usually represented as *a* and *ə*; while other non-high vowels sometimes occur, they do not ordinarily contrast with *a* and *ə* phonemically. We will see soon that a three-way contrast of non-high vowels is one of the features of the Central dialects.

Apical and retroflex vowels are a feature found in many Chinese dialects, but are especially prevalent in the Northern and Central groups. In addition to the unrounded apical [ɿ] and the retroflex [ʅ] vowels described for the standard language, a few dialects also possess rounded counterparts, a rounded apical vowel [ʉ] and a rounded retroflex vowel [ɥ] respectively.

Dialects of the Northern zone are generally rich in diphthongs and triphthongs, in conspicuous contrast to the Wú dialects, for example, where such combinations are much rarer. This may be viewed as a conservative feature when compared with Middle Chinese, in which vocalic clusters abounded.

From a comparative point of view, Mandarin dialects have a small number of tones. A few dialects have only three tones, such as Línchéng in Héběi. More typically four tones are found, as in the Peking dialect. In the Jiāng-Huái region five tones are not uncommon. A few Mandarin dialects with six tones have been reported along the boundary that separates the Northern and the Central dialects. Many Mandarin dialects (perhaps a majority) show a tonal pattern similar to that described for Peking: the *píng* tone occurs in two registers, the *yīnpíng* in words which in Middle Chinese had voiceless initials, and the *yángpíng* in words which in Middle Chinese had voiced initials: here and there one can encounter dialects in which these two categories have merged, for example, Zhāngjiākǒu in Héběi and Tàiyuán in Shānxī. No Mandarin dialect (as far as is known) shows a register distinction for the *shǎng* tone; words with voiced obstruent initials of this

Figure 7. Norman, Jerry. 1988. p. 194.

References

- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Karlgren, Bernhard. 1915-1926. *Etudes sur la Phonologie Chinoise*. Archives D'etudes Orientales, in 4 volumes, Vol. 15:1-4. Published by J.-A. Lundell. Brill, Leiden/Uppsala.
- Linguistics Section, Department of Chinese Language and Literature, Beijing University (北京大学中国语言文学系语言学教研室) (ed.). 1989. *Hanyu Fangyin Zihui* (汉语方音字汇). [Phonetic Dictionary of Chinese Dialects]. Second edition. Beijing: Wenzhi Gaiye Publishing House [文字改革出版社]. (The first edition was published in 1962.)
- Norman, Jerry. 1988. *Chinese*. Cambridge Language Surveys series. Cambridge, UK: Cambridge University Press.
- The Unicode Consortium. 2007. *The Unicode Standard*, Version 5.0.0, defined by: The Unicode Standard, Version 5.0. Boston, MA: Addison-Wesley.
- Yuan, Jiahua (袁家骅) et.al. 1989. *Hanyu Fangyan Gaiyao* (汉语方言概要) [Outline of Chinese Dialects]. Second edition. Beijing: Yuwen Publishing House [语文出版社]. (The first edition was published in 1960.)