Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type:	**Working Group Document**
Title:	**Proposal to add six phonetic characters for Scots to the UCS**
Source:	**Michael Everson**
Status:	**Individual Contribution**
Date:	**2019-03-23**

This proposal requests the encoding of two combining characters and four spacing characters used in linguistic transcriptions of Scots. If this proposal is accepted, the following characters will exist:

1AC1   COMBINING LATIN SMALL LETTER W BELOW

       • used in transcriptions of Scots to indicate a voiced labialization

1AC2   COMBINING LATIN SMALL LETTER TURNED W BELOW

       • used in transcriptions of Scots to indicate a voiceless labialization

AB68   LATIN SMALL LETTER TURNED R WITH MIDDLE TILDE

       • used in transcriptions of Scots to indicate a velarized dento-alveolar approximant

AB69   MODIFIER LETTER SMALL TURNED W

       • used in transcriptions of Scots to indicate a voiceless bilabial realization

AB6A   MODIFIER LETTER LEFT TACK

       • used in transcriptions of Scots to indicate a fronted vocalic realization

       x 02D4 modifier letter up tack

       x 0318 combining left tack below

AB6B   MODIFIER LETTER RIGHT TACK

       • used in transcriptions of Scots to indicate a backed vocalic realization

       x 0319 combining right tack below

The characters here are used in the *Scottish National Dictionary* and the *Linguistic Atlas of Scotland (Scots Section)*, which are, as can be expected, focused on the specific sounds found in Scots dialects. Some aspects of the transcription in those sources differ from those in standard IPA, and the missing characters are proposed for encoding here.

1. Unicode Character Properties. Character properties are proposed here.

```
1AC1;COMBINING LATIN SMALL LETTER W BELOW;Mn;220;NSM;;;;;N;;;;;
1AC2;COMBINING LATIN SMALL LETTER TURNED W BELOW;Mn;220;NSM;;;;;N;;;;;
```

```
AB68;LATIN SMALL LETTER TURNED R WITH MIDDLE TILDE;Ll;0;L;;;;;N;;;;;
AB69;MODIFIER LETTER SMALL TURNED W;Lm;0;L;<super> 028D;;;;N;;;;;
AB6A;MODIFIER LETTER LEFT TACK;Sk;0;ON;;;;;N;;;;;
AB6B;MODIFIER LETTER RIGHT TACK;Sk;0;ON;;;;;N;;;;;
```

2. Bibliography

Grant, William, ed. [1931]. *The Scottish National Dictionary*. Volume I: A–Bitteraks. Edinburgh: Scottish National Dictionary Association.

Grant, William, and David D. Murison, eds. 1965. *The Scottish National Dictionary*. Volume VI: Langsome–O. Edinburgh: Scottish National Dictionary Association.

Mather, J. Y., and H. H. Speitel, eds. 1986. *The Linguistic Atlas of Scotland*. (Scots Section: Volume III: Phonology.) London, Sydney, and Dover: Croom Helm. ISBN 0-85664-716-0.

3. Figures.

Inventory of Principal Consonant Symbols

	stop	fricative	nasal	lateral	continuant	roll/flap
bilabial	p b	ɸ β	m		w ʍ	
labio-dental		f v				
dento-alveolar	t d	θ ð s z	n	l	ɹ	r ɾ
retroflex	tʳ ḍ	ṣ	ɳ	lʳ		ʈ
palato-alveolar	ţ ḍ	ɕ ʑ ʃ ʒ				
palatal	c ɟ	ç j	ɲ			
velar	k g	x ɣ	ŋ			
uvular	q G	χ ʁ				ʀ
glottal	ʔ	h				

Notes

(1) ₊ denotes *palatalisation*, e.g. s₊, z₊, l₊, n₊, k₊, g₊, x₊; but ţ, ḍ.

(2) *Retroflexion* is indicated in the symbols ḍ, ɳ, ɾ, ṣ, but also by use of superior 'r' in the case of t, l, i.e. tʳ, lʳ; and also where vowels are retroflex or 'r-coloured', e.g. ɜʳ, əʳ.

(3) In ɫ, ɴ, ʀ, ɨ there is *velarisation*; in x̫, x̫, ɸ̫, ɸ̫, there is *labialisation*. ɫ has lateral friction.

(4) ḍ, ɾ̥, ɹ̥, ʀ̥ show *devoicing*; ḍ, ţ, ɾ, ɾ are *dental*.

(5) ˘ is used in oĕ and ŏe, where oĕ is a *falling* diphthong and ŏe is a *rising* diphthong.

General

(1) Brackets: [] *phonetic*, / / *phonemic*, \ \ *polyphonemic*.

(2) #, the symbol following is word-initial.

(3) ≡, as in [ä⊥] ≡ /a/, should be interpreted '[ä⊥] is an allophone of /a/'.

Figure 1. Example from Mather and Speitel 1986:xviii showing ɨ LATIN SMALL LETTER R WITH MIDDLE TILDE, ◌̫ COMBINING LATIN SMALL LETTER W BELOW, and ◌̫ COMBINING LATIN SMALL LETTER TURNED W BELOW. Also shown is a peculiar glyph for the voiceless uvular stop /q/.

Phonetic Symbols and Diacritics

The phonetic symbols and diacritics used in the tables and lists which follow are, in the main, those of the IPA. However, we have adopted, in some cases, particular uses and arrangements of these. The final inventory is the result of some rationalisation and standardisation of the symbols used by field-workers. Finer details and discriminations are reflected by use of diacritics.

Inventory of Principal Vowel Symbols

	FRONT		CENTRAL		BACK		
HIGH	i	y	ɨ	ʊ	ɯ	u	
	e	ø	ɪɫ	ɵ	ɤ	o	MID
LOW	ɛ	œ	ə		ʌ	ɔ	
	a		ɐ		ɑ	ɒ	
	non-round	round	non-round	round	non-round	round	

Notes

(1) · denotes length; : is used (rarely) where there is a second contrastive length.

(2) ⊢ indicates *backer*, ⊣ *fronter*, ⊥ *higher*, ⊤ *lower*; these marks are placed after the vowel symbol, e.g. a⊢, u⊣, e⊥, ɛ⊤. Note that where the vowel is long, the length mark comes first, immediately following the vowel symbol, so a·⊢, u·⊣, e·⊥, ɛ·⊤.

(3) ¨ signifies *centralised*, e.g. ë, ö, ÿ, ö.

(4) ʊ has been used instead of ɷ.

(5) In the *tables only* * indicates that the vowel phoneme was elicited in *one* word, e.g. ʌu*. (Elsewhere it may signify a footnote.)

Figure 2. Example from Mather and Speitel 1986:xviii showing ⊣ MODIFIER LETTER LEFT TACK and ⊢ MODIFIER LETTER RIGHT TACK.

	Description	Abbrev.	
ʃ	breathed after-blade fricative	b.a-bl.fric.	Sc. *shoogle* ; Eng. *shut* ; Middle Sc. *schipp*.
tʃ			⎰Sc. *chape*, mu*tch* ; Eng. *ch*eap, ⎱ ri*ch*, hi*tch*.
ʒ	voiced after-blade fricative	v.a-bl.fric.	pu*s*ion (poison).
dʒ			⎰Sc. *jouk*, *fugie*, *gimp* ; Eng. ⎱ bri*dge*, *jig*, *gin*.
x	breathed back fricative	b.bk.fric.	Sc. lo*ch* ; Ger. la*ch*.
xʍ	breathed back fricative rounded	b.bk.fric.r.	With rounded lips. s.Sc. lau*ch*, sou*cht* ; Ger. au*ch*.
ʍ	breathed lips-back fricative	b.l-bk.fric.	Sc. *wh*aup ; Eng. *wh*isper.
ʔ	breathed glottal plosive or stop	b.gl.pl.	Heard in Central Sc. dialects before or as a substitute for *p, t, k*.

Point means that the place of formation is between the point of the tongue and the apex of the upper gum.

Front means that the place of formation is between the middle of the tongue and the middle of the hard palate.

Back means that the place of formation is between the back of the tongue and the back palate.

Glottal means that the place of formation is between the vowel chords in the larynx.

Figure 3. Example from Grant [1931]:xliii showing ʍ MODIFIER LETTER SMALL TURNED W.

LAUCH, *v., n.*[1] Sc. forms and usages of Eng. *laugh.* For I.Sc. forms see LAICH, *v.*
I. *v.* **A.** Forms. Pr.t.: *lauch* (Gen.Sc.); *lach*; *laach* (Sh. 1952 Robertson and Graham *Sh. Dial.* 33); *liach* (Ork. 1909 *Old-Lore Misc.* II. i. 30); ¶*leugh* (Lnk. 1890 H. Muir *Rutherglen* 12). [lǫ(:)x, lɑ(:)x; s.Sc. + ‡lɑxʍ.) See P.L.D. § 111.]

Figure 4. Example from Grant and Murison 1965:11 showing ʍ MODIFIER LETTER SMALL TURNED W.
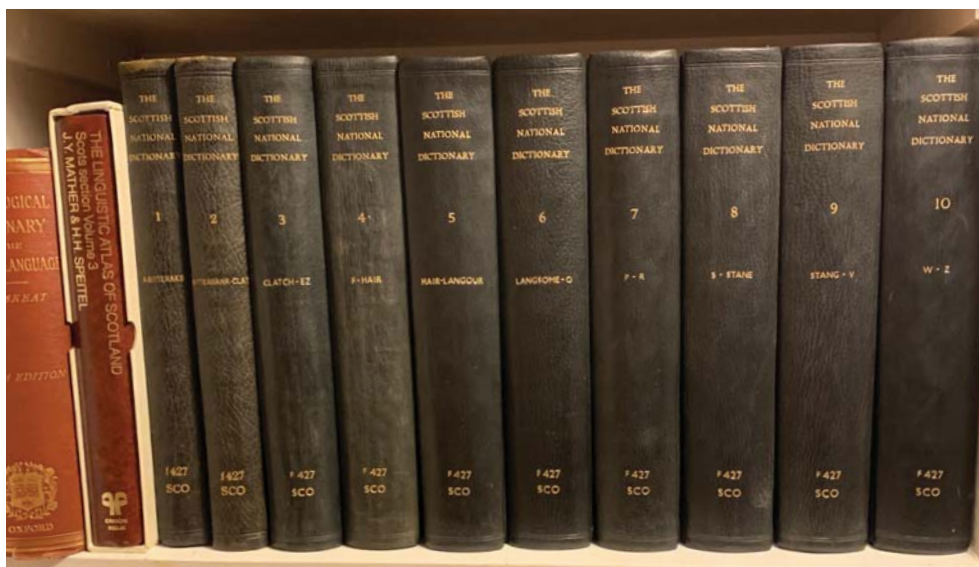

Figure 5. The Scottish National Dictionary, 1931–1976.

A. Administrative

1. Title
Proposal to add six phonetic characters for Scots to the UCS
2. Requester's name
Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2019-03-23
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
Yes.
6b. More information will be provided later
No.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
No.
1b. Proposed name of script
1c. The proposal is for addition of character(s) to an existing block
Yes
1d. Name of the existing block
Combining Diacritical Marks Extended (2); Latin Extended-E (4)
2. Number of characters in proposal
6.
3. Proposed category (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)
Category A.
4a. Is a repertoire including character names provided?
Yes.
4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
4c. Are the character shapes attached in a legible form suitable for review?
Yes.
5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson.
5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes.
6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
7. 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.
8. 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/ Public/UNIDATA/UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.
See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
No.
2a. 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.

2b. If YES, with whom?

2c. 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?

Germanicists, Anglicists, dialectologists, lexicographers, and Scots.

4a. The context of use for the proposed characters (type of use; common or rare)

Common in Scots linguistic reference material.

4b. Reference

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

Yes.

5b. If YES, where?

Various publications.

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

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Accordance with the Roadmap. Keep with other Latin characters.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

No.

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

Yes.

10b. If YES, is a rationale for its inclusion provided?

Yes.

10c. If YES, reference

Cross references point to the related but different characters.

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

Yes.

11b. If YES, is a rationale for such use provided?

No.

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?