Title: **Evolution of the Tifinagh script in Unicode**
Source: **Paul Anderson**
Status: **Individual Contribution**
Action: **For consideration by UTC**
Date: **02-Aug-2010**

This document aims to outline some of the problems facing the Tifinagh script in Unicode and to suggest possible future directions.

The Berber language group encompasses many language variants spoken by communities across a huge geographical range, with lifestyles from settled to nomadic, living in cities, desert oases and deep desert. For convenient reference, I group the major language variants together into 'Northern Berber' and 'Tuareg'.

Those language variants with a written tradition are written in one or more scripts - Latin, Arabic and Tifinagh. In some cases there is an official script for a language variant – Latin in Niger and Mali for Tuareg variants; Tifinagh in Morocco and Latin (de facto) in Algeria for Northern Berber variants.

There are different styles of the Tifinagh script in use according to regional tradition or to the influential modern example adopted by writers or authorities. The styles, all in current use, can be grouped into traditional Tuareg Tifinagh, modern vowelled Tuareg Tifinagh, and neo-Tifinagh (used for Northern Berber). Even within each of these style groups there are variations - in how vowels are marked, regional letter repertoire and typographic convention.
See the two tables 1a and 1b * here:
http://www.win.tue.nl/~aeb/natlang/berber/tifinagh/tifinagh-mondeberbere.html
or the same tables here:
http://www.mondeberbere.com/langue/tifinagh/

Traditional Tuareg Tifinagh has many consonant letters composed of dots, very limited vowel marking, and often biconsonant ligatures. Knowledge of this style is still passed down the generations among the Tuareg.
Some major regional styles in this group are shown in the first table, 1a, in columns H, G, D, Y, W and T.

Modern vowelled Tuareg Tifinagh styles each add some novel system of Tuareg vowels to regional consonant letters of traditional Tifinagh from table 1a. An example from Niger is shown in table 5. I will examine these styles in more detail later.

Neo-Tifinagh refers to modern efforts to create an improved repertoire (motivated by writing Northern Berber). The first table, 1a, shows the repertoires used as a source of letters on the left – the traditional Tuareg Tifinagh forms discussed above, along with the ancient scripts: western Libyic (L/oc), eastern Libyic (L/or) and Saharan (Sah).

The first table, 1a, (except for column AB) is attributed to Prasse (1972), so it is probably assembled from the same Tifinagh tables of Prasse's 1972 "Manuel de grammaire touarègue" that are referenced in the original Unicode Tifinagh proposal (1) .

Neo-Tifinagh introduced innovations for writing vowels (the same motivation as modern vowelled Tuareg Tifinagh), and for simplifying and clarifying typography, for example preferring letters not composed of dots. Column AB shows the resulting neo-Tifinagh repertoire of the Academie Berbère. Table 2a shows some other neo-Tifinagh repertoires - IRCAM's alphabet is another (not shown). All are broadly similar in composition and scope.

When the Tifinagh script was first accepted into Unicode, the majority of the letters were taken from IRCAM (the national governmental Berber body of Morocco). IRCAM had arrived at a particular repertoire of neo-Tifinagh letter forms comprising an essential basic subset and an extended subset suitable for more phonetic writing. IRCAM's basic subset is the only repertoire that a nation has declared official. The remaining letters entering Unicode consisted of attested Tuareg letters and a few attested non-IRCAM neo-Tifinagh.

At the moment, the Tifinagh script in Unicode has not been simplified to a common character set, despite having a subset of characters theoretically perfectly suited to writing both Northern Berber variants and Tuareg language variants using different font renderings to achieve suitable Tifinagh styles. Neither does it have an adequately large character set for writing both neo-Tifinagh script and Tuareg traditional Tifinagh script with the same font, because of incompatible typographic conventions, and not all regional letter variants are covered.

Two possible responses would be:

– to add further characters as required for unsupported regional letter repertoires; perhaps also to add Tuareg versions of letters such as YAN/YAL/YA/YU which have different typographical conventions to the equivalent neo-Tifinagh letters. With the presence of Tuareg forms of letters guaranteed for all fonts, font support for the biconsonant ligatures of TIFINAGH CONSONANT JOINER could be mandated.
– to unify on a common Tifinagh subset (a set corresponding in scope to those characters with 'unmarked' names i.e. (TIFINAGH LETTER <X>, not e.g. TIFINAGH LETTER TUAREG <X>).

The following sections will explore the background and analyse the implications of these responses.

The language variants written in Tifinagh

I use the term "language variant" rather than "dialect" or "language" to simplify the discussion, though sociolinguistically and phonologically the major language divide would seem to fall between the forms spoken in the North and the Tuareg forms with two extra long vowels, an extra short vowel, and stress contrasts between verb forms.

Tuareg language variants are highly mutually intelligible, and Tuareg can be viewed as a macrolanguage (which has the ISO-639-3 code 'tmh').

Northern varieties (grouped here as 'Northern Berber') are more divergent because of long geographical isolation from each other historically, but still show a striking fundamental linguistic unity on multiple levels. IRCAM are working on formalizing a converged language from the Moroccan variants, which is taught in schools at later stages, but the Algerian language variants have not yet been incorporated into it.

I have not considered in detail the oasis variants, or Zenaga in the extreme West, or Siwi in the extreme East. Zenaga is almost extinct, Siwi is less-researched and has not been written in Tifinagh except experimentally, and the oasis variants either fall among the Northern variants, have no recent history of Tifinagh use, or have not been prominent in modern printed Berber literature.

The different styles (letter repertoires plus typography) of Tifinagh

These styles all perform the essentially same function, allowing for differences in pronunciation and vowel repertoire between language variants, and they overlap greatly in letters and typographical techniques used. I group them separately below only to help explain their features. They vary in the amount of innovation in their letter forms and typography and in how extensive their vowel systems are.

Traditional Tuareg Tifinagh

This style is still widely known (and therefore often used as the basis for literacy programmes) but rarely used other than for short inscriptions and letters. See (2) for examples of writing.

In table 1a, the traditional Tuareg repertoires are labelled according to the name of the region, the local Tuareg confederation or its dialect. The columns are H (Hoggar/Ahaggar, Algeria), G (Ghat, Libya), D (Adrar/Adghagh – Adrar n Ifoghas, Mali), Y (Ayer/Aïr, Niger), W (Iwəlləmədan; Azawagh/Azawak/Azaouad/Azawad;Tawellemet, Niger-Mali), T (Tanslemt-dialect of Timbuktu dialects, Mali).
Table 2a below has most of the same repertoires (I have added the corresponding letter codes from table 1a) with minor differences – this table was featured in the original Tifinagh proposal (1).

Systems vary in which vowels are written, if any, and in which positions. No version has letters for the long vowels 'e' or 'o', or a way to indicate the short 'a' vowel or contrastive stress.
In the absence of vowel marking, traditional Tifinagh often uses biconsonant ligatures to show when no vowel is present between two consonants. There are tables of ligatures and example texts in (1) and (2). Here are some examples for 'rt' from the Ahaggar region (1):-

Between regions, some letters are the same and others differ because of their long history of use over a wide area. Sometimes letters are the same but have come to denote different sounds, probably because of dialectal sound changes. It is possible that some regional variations have not yet been documented – some language variants have been better studied than others, the language has generally been more of a priority than Tifinagh, and political instability has obstructed study.

Letters consisting of vertical strokes such as YAN (|) or YAL (||), when written in clusters, would cause visual ambiguity ('nn' would be '||', for example) so these letters are leaned and offset in various ways where necessary (fig. 4). See (1), section D9, for details of the different conventions in use.

Many consonant letters consist of dots, which are always written filled, not as open circles.
Examples are ⁖ ⁚ ⠇ ⁞ ⁛

TABLEAU DES TIFINAGH
(alphabet touareg)

	Ahaggar (Algérie) H	Ghat (Libye) G	Aïr (Niger) Y	Azawagh (Niger-Mali) W	Adghagh (Mali) D
a	·	·	·	·	·
b	ⵔ	ⵔ	ⵔ	ⵔ	ⵔ
d	ⵎ	ⵎ	ⴻ	ⴻ	ⴸ
ḍ	ⴻ	ⴻ			ⴻ
f	ⵊⵎ	ⵊⵎ	ⵊⵎ	ⵊⵎ	ⵎ
g	ⵝ	·⫲·	⫲	⫲	·⫲·
gʸ	⫲	ⵊ			ⵊ
ɣ	ⵗ	ⵗ	ⵗ	ⵗ	ⵗ
h	ⵞ		ⵌ	ⵌ	ⵞ
j			ⵣ	ⵣ	ⵣ
k	ⵗ	ⵗ	ⵗ	ⵗ	ⵗ
l	‖	‖	‖	‖	‖
m	ⵎ	ⵎ	ⵎ	ⵎ	ⵎ
n	ⵉ	ⵉ	ⵉ	ⵉ	ⵉ
nʸ	≠				
q	···	···	ⵆ	ⵗ	···
r	ⵔ	ⵔ	ⵔ	ⵔ	ⵔ
s	ⵙ	ⵙ	ⵙ	ⵙ	ⵙ
š	ⵯ	ⵯ	ⵯ	ⵯ	�8
t	+	+	+	+	+
ṭ	⵴				
w	ⵗ	ⵗ	ⵗ	ⵗ	ⵗ
x	ⵗⵗ	ⵗⵗ	ⵗⵗ	···	ⵗⵗ
y	ⵠ	ⵠ	ⵠ	ⵠ	ⵠⵠ
z	ⵥ	ⵣ	ⵥ	ⵥ	ⵞⵠ
ẓ	ⵣ	ⵞ		ⵗ	ⵣ

Table 2a – Mohammed Aghali-Zakara 1993. See also (1).

Modern vowelled Tuareg Tifinagh

There have been several efforts, both by associations and individual artist-activists. These repertoires build on the traditional Tuareg Tifinagh forms rather than neo-Tifinagh (and therefore still need context-dependent letter leaning), adding new letters only for vowels. The different vowel systems are sketched out in (1). Some are discussed in (3).

One such repertoire is from the UNESCO-aided Association for the Promotion of Tifinagh (=APT) in Agadez, Niger. See table 5 for the letter repertoire (based mainly on the traditional Aïr letters from table 2a) and fig. 4 for an example of text. Although the letter forms for 'i', 'u' and schwa ("e muet") differ from IRCAM's (see table 6), the system of vowel writing is essentially the same, extended with the two Tuareg long vowels 'e' and 'o' ('<' and '–'), with a breve on YA (ӗ) for the short 'a' vowel, and with the use of circumflexes to indicate contrastive stress.

Classes organized by the APT in 2003 taught approximately two thousand Tuareg the modernised script (of course, the traditional Aïr repertoire on which it is based is widely known). The APT printed and distributed several publications in this style.

Another repertoire is from SIL, based on the traditional Tawellemet letters, with a slightly different vowel system from the APT – see (1) and Unicode internal document L2/10-147.

Rissa Ixa's (a Tuareg artist/calligrapher) and Hawad's (a Tuareg poet/calligrapher) systems are discussed in (3). In reference 3 (published in 1993) its author describes a questionnaire about people's Tifinagh knowledge and use. 100 copies were distributed (in Niger, it seems from the document) and 60 were returned. Five respondents claimed familiarity with Hawad's vowel system and one had adopted one of Rissa Ixa's vowel letters (÷). I have no further details about the methodology of the questionnaire's distribution.

Others vowel systems are sketched out in (1) but I have no details about current use.

Northern neo-Tifinagh

Different neo-Tifinagh repertoires were created at different times, by different people, in different regions. They overlap greatly, each repertoire having taken letters in different proportions from the available sets – Tuareg Tifinagh, ancient Saharan, and ancient Western and Eastern Libyic – and adapted them with typographical innovations to a greater or lesser extent. The Internet and printed publications have allowed the repertoires to be widely discussed, so which form is seen in a particular region therefore depends on awareness of and need for developments and on which repertoire has recently found general concensus, rather than on deep local tradition.

Neo-Tifinagh (i.e. modernised) repertoires avoid punctiform letters because illegible consonant sequences are possible – see (4), Aïcha Bouhjar, "Le système graphique Tifinaghe-Ircam", p.54. Also, in IRCAM neo-Tifinagh, for vowel letters (where dots are retained), the dots are often written as open circles to avoid confusion with punctuation. For example TIFINAGH LETTER YU = Tuareg yaw in Tuareg usage appears as "•" but in IRCAM neo-Tifinagh it is "○". In another IRCAM innovation, the constituent vertical lines of YAL (ⵀ) are connected with a small link to avoid the need for context-dependent leaning - see figure 4 and (1) section D9. This is also done for YAF (ⴻ, instead of ⵉⵏ), presumably to avoid confusion with YAM (ⵏ) or square brackets.

IRCAM's repertoire is a solid official standard in Morocco (other letters are still seen in Morocco, for example for writing Rifian more phonetically than the national standard specifies). See (4), "Le système graphique Tifinaghe-Ircam" for a full explanation of its development. In Algeria however you see many street and building signs (such as bank logos), posters, T-shirts, political party and association business cards, graffiti examples and titles of publications written in diverse neo-Tifinagh varieties (occasionally with very idiosyncratic letters taken from ancient Libyic or Saharan) that reflect the time that the text was written and the knowledge of its author.

Although fewer people now are familiar with Tifinagh in Algeria than a few years ago, because of the dominance of the Latin script for Kabyle Berber literature (Tifinagh's role tends to be rather a cultural icon), a concensus has emerged there, and it differs somewhat from IRCAM's standard (see figs. 3a and 3b). The hollow dot and connected line innovations of IRCAM are lacking, so context-dependent letter leaning is still needed, and the letter forms in the two systems for the same sounds sometimes differ. Also, since the sounds 'tch' and 'dzh' are commoner and more fundamental in Algerian Berber language variants than in Morocco, they are written with a single dedicated letter each:

ⵣ (not to be confused with TIFINAGH LETTER YAP "ⴆ") and ⵝ.

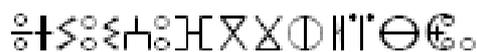See (4), "Les caractéristiques phoniques de l'Alphabet Tifinage-Ircam" and (1) section D2.2.

ⵣ⵰ⵜⵙ⵾ⵞ⵾ⴽⵀⵂ⵼ⵌⵝⵝⴹⵏⵏⵏⴻⴲ⵰

Figure 3a - Hapax Berbère font letters (IRCAM neo-Tifinagh forms)

÷ⵜⴳⵧ⵾Σ⵰⵰ⵌⵙⵝⵝⵧⵁⵏⵝⴹ⵾ⵦ

Figure 3b – Equivalent Algerian neo-Tifinagh forms to 3a

Figure 4 (Tišiway n Ayər, APT Agadez 2006 p. 20).
TIFINAGH LETTER YE '<' and YO '–' are circled,
along with context-dependent leaning of
YAN (line 6) and YAL (line 5)

Table 5 (APT Niger modern vowelled Tuareg Tifinagh repertoire, from
"Lexique illustré tamajaq-français", 2006)

	APT	IRCAM
"i"	ⵚ - TIFINAGH LETTER YAY (Consonant Y is written with YI)	ⵉ - TIFINAGH LETTER YI (commonly written in Tuareg for consonant Y)
"u"	‖ - Consonant W is written with YU (Tuareg YAW) which takes the form =	ⵓ - TIFINAGH LETTER YU (Tuareg YAW)
"ə" - schwa	⊃	ⵛ – TIFINAGH LETTER YEY

Table 6 (Differences between APT vowel letter forms and IRCAM's)

Brief history of Tifinagh in Unicode

When the Tifinagh letters were first added to the UCS, the main resource used was the vowelled
neo-Tifinagh repertoire developed by Morocco's IRCAM for Moroccan variants of Northern Berber.
IRCAM designated a basic subset of the repertoire for writing in a standard Tifinagh orthography,
and when Tifinagh was declared as the official script for these language variants by Morocco, this
subset became the official alphabet.

Another IRCAM subset, called the 'extended' set, consisted of other neo-Tifinagh letters that extended the basic set without duplication of sound, for example U+2D36 TIFINAGH LETTER YAJ, U+2D32 TIFINAGH LETTER YABH. These letters were not needed by IRCAM in the basic set because the letters are used for writing regional or foreign features phonetically.

In Unicode, the IRCAM letters (basic and extended), plus some other neo-Tifinagh letters that did not duplicate sounds from the IRCAM sets, were added simply as "TIFINAGH LETTER <X>" where <X> corresponds to the sound represented by the letter.

Other attested non-IRCAM neo-Tifinagh and Tuareg variant letters were added to Unicode, with names such as U+2D4C TIFINAGH LETTER TUAREG YAZH. Consideration of any letters for which the situation was not yet clear was deferred, though most letters already identified are available in the PUA of the Hapax Berbère reference font (5).

Later, a Tifinagh punctuation character called "Tazarast" was added, along with a Tifinagh joiner character to indicate biconsonant ligatures in traditional Tuareg writing (2).

Writing the different Tifinagh styles in Unicode

Here I will discuss how effectively the different styles of Tifinagh can be written with the current Unicode set, using variant letters where needed for a particular style - the conventional way to obtain the correct visual letter form.

Neo-Tifinagh

Writing in standard IRCAM style is of course easy, since the necessary characters were encoded completely from the start. This is the only style that uses the basic IRCAM letter subset exclusively.

For Algerian neo-Tifinagh, only some of the letters that differ from equivalent IRCAM ones are encoded, for example TIFINAGH LETTER BERBER ACADEMY YAH and BERBER ACADEMY YAJ. One missing letter can be written as YA with a circumflex (⵿). Others are not encoded, such as the letters for 'y' (an inverted ⎍ YAW, see fig. 3b) and for 'tch': ⵣ (similar to YAP).

Currently, writing Algerian-style is done with non-Unicode fonts that map corresponding Latin letters to the neo-Tifinagh forms. Symbols and precomposed letters with diacritics available on the French AZERTY keyboard are also used. See: http://www.mondeberbere.com/langue/polices.htm for the "Afus deg wfus" fonts. Such fonts are also distributed by the national Berber High Commission (HCA).

Algerian writing has vowel letters written with filled dots, so any fonts that show vowel letter dots as hollow would look strange. It requires font support for traditional Tuareg-style context-dependent letter leaning, as the constituent vertical lines of YAL (⵫) are not linked unlike IRCAM's YAL (ⵣ).

Traditional and modern vowelled Tuareg Tifinagh

Most traditional consonant letters are in the UCS; some that are not could be considered visual variants of existing letters. Others are not encoded but may not be important. See table 1a.

Tuareg styles mix variant letters with the basic IRCAM set, their traditional letters often differing from IRCAM's modern ones for the same sounds.

In some cases this mixing of the basic IRCAM set and variant letters for Tuareg causes difficulties, because while neo-Tifinagh fonts can arguably choose for the basic set whether to follow innovations such as linked bars for YAL and YAN, and hollow dots for YA and YU, Tuareg fonts can not:
- The hollow dots for vowels are incorrect for Tuareg and also do not match the variant consonants formed with filled dots, such as Tuareg YAK. However, the punctiform consonants could be written using hollow dots to match the vowels, and the letters would be understood - but they would look strange.
- Contextual shaping works differently for combinations of the letters TIFINAGH LETTER YAL and TIFINAGH LETTER YAN when writing neo-Tifinagh and when writing Tuareg (and even differs between traditional and APT/SIL modern vowelled Tuareg orthographies – see (1) section D2). In IRCAM neo-Tifinagh the vertical lines within a letter are linked, so no confusion can arise when letters are clustered, but in Tuareg usage the lines must be 'leaned' to show which letter they belong to (fig. 2). Even for neo-Tifinagh the choice of using linked bars for YAL and YAN is not strictly free for a font, since Algerian-style neo-Tifinagh for Northern Berber also uses letter leaning and IRCAM practice would look strange.

Writing Tuareg is only currently possible in Unicode in the traditional style without vowels (or with very limited vowel marking). There is a Unicode proposal in progress to address this in principle by adding TIFINAGH LETTER YE and TIFINAGH LETTER YO from the APT Niger system of Tuareg vowelling, but these would have to be used in conjunction with non-APT consonant letters and vowel letter equivalents because of other APT letters not yet encoded. For APT Niger-style modern vowelled Tuareg Tifinagh, letters are missing from Unicode (schwa - Ɔ, u - ıı, tch - Ⲧ), though there are existing equivalents differing in letter form (YEY - ⵚ, YU – ⵛ used as YAW by APT, YACH - ⵞ).

In the absence of vowelling, the traditional ligatures (2) needed for marking consonant pairs with no vowel between are as yet only rarely even partially implemented in fonts. (Even then, they are not for use with the TIFINAGH CONSONANT JOINER but for the earlier mechanism of ZERO WIDTH JOINER). The required mapping table is large because of the number of variant Tuareg letters, and ligatures might even vary by region. Also, if the font used is neo-Tifinagh, neo-Tifinagh YAF alone could be read as the traditional 'nf' ligature (ⵂ versus ⵃ).

Speculative uses

Outside the IRCAM basic set there are already extra letters likely to be sufficient for representing the richer phonology of Zenaga and any similar oasis language variants. Siwi seems to have 'e' and 'o' which could be written as for Tuareg. Though Zenaga is near-extinct, Siwi is apparently thriving, and although it is not typically written in Tifinagh, local awareness of the possibility exists (see http://www.haberlah.com/hosted/siwa/siwa_language.html).

For writing dialect dialogue phonetically, or for approximating foreign words, such extra letters are likely to cover most needs, perhaps in conjunction with diacritics especially for vowels.

Overall

Although Unicode encodes many variant Tifinagh letters, these do not cover all those needed for different attested styles of Tifinagh writing. (However, as I will discuss shortly, using a Tifinagh character subset shared by all styles, leaving the graphical variation to the fonts, would theoretically cover the orthographic requirements).

Even where the existing letters are sufficient, although the variant letters are encoded explicitly, the different ways of writing Tifinagh still cannot all be shown correctly by any one font, because of differences in the customary appearance of the letter and differing solutions for sequences of letters with vertical lines. This situation is rather like for traditional and simplified Chinese.

An arbitrary Unicode font would be unlikely to render the diacritics attractively for vowelled Tuareg, though it could handle biconsonant joining simply by showing some invariant symbol instead of using a ligature table (as suggested in (2)).

Since the Tifinagh character mappings from the IRCAM set to the variant repertoires are one to one, and different fonts could capture the variety, the original reason for encoding the variant letters individually seems to have been the greatly differing glyphs, or to allow different Tifinagh styles to be written side-by-side by the same font (not actually possible exactly, as we have seen).

Possible responses and implications

Full unification

As explained, there are many different versions of the Tifinagh alphabet in use in different regions, both in the North and for Tuareg. However, not counting the extra vowels of Tuareg, where different systems have been suggested, it is possible to tabulate all these versions side by side, correlated by sound – and historically this is what all researchers have intuitively done.
See tables in http://www.win.tue.nl/~aeb/natlang/berber/tifinagh/tifinagh-mondeberbere.html and table 2a.

There are some gaps depending on the phonology of the language variants, and some letters are used in different areas to represent different sounds because of historical sound shifts, but in general the unity of the script is easily visible. Each region differs only in a handful of positions from the others, and each region requires a repertoire of letters to represent almost the same set of sounds. In each case the same logical letter is intended, for example TIFINAGH LETTER YAT. There is therefore a clear shared superset of letters required corresponding in scope to the Unicode Tifinagh with 'unmarked' names, and the script could be fully unified on this basis. There is no need for displaying different regional repertoires side by side in the same font (doing this for neo-Tifinagh and traditional Tuareg styles is impossible with the current approach anyway, as already explained in the preceding section). Two language variants in the same document would either use the same Tifinagh writing style (e.g. on the computer command line with the system font) or different fonts as appropriate.

The fact that the character naming in Unicode for Tifinagh forms a fundamental alphabet with unmarked names of the form TIFINAGH LETTER <X>, plus variant letters for the same sounds with more specific names, seems to indicate that the original encoders of the script were already thinking in these terms, though they decided to begin encoding the variants individually (without capturing enough of them initially to write all the different repertoires).

There is one exception to this character equivalence between repertoires – APT Niger Tuareg encodes emphatic letters as the unemphatic letter plus combining dot below, for example YAT plus COMBINING DOT BELOW instead of standalone letter YATT. That poses no problem for unification of the script, however.

Tuareg Tifinagh Unicode letters do not yet have widespread use, probably partly because of the lack of font support for biconsonant ligatures and the lack of letters for Tuareg-specific vowels in Unicode. There does seem to be some pre-Unicode font writing around. At the same time, the IRCAM alphabet, which has taken root in Morocco, would not be disturbed by full unification. In Algeria, there is not yet an official script and the de facto standard for general text is Latin script, so although non-Unicode fonts often have to be used to write Tifinagh for poster titles etc., a push to encode the neo-Tifinagh used in Algeria is unlikely. This all offers a clear opportunity.

Berbers are highly independently-minded, and in view of their position as minority populations, governmental regulation of the appearance of their traditional repertoires is likely to be resented. By contrast, unification of the character set offers simplicity, and allows sharing of tools such as keyboards, a huge advantage where the contributor pools for IT support for individual language variants are small. See the unified Unicode Berber/Tuareg keyboard and corresponding regional fonts at http://www.akufi.org/tools/index.html

Repertoires as yet unsupported by Unicode or even undocumented can be supported quickly, and fonts can ensure that typographic features for their given repertoires are consistent. For example, a font for a style that requires a particular way of letter leaning can implement it correctly, and a font with punctiform consonants would avoid vowels with hollow dots. Ligatures that are mandatory for traditional Tuareg could be optional and cosmetic in fonts for Northern Berber. Punctuation such as comma and hyphen would look different in APT Niger-style fonts.

This approach leaves considerable latitude to font designers for regional and calligraphic variation in the visual letter forms. Mixed repertoire (and even historical repertoire) fonts are possible, and choice of font is easy for a text, allowing use of the Tifinagh script to evolve freely as useful features or popular letter forms are adopted.

There would probably be no need to extend a unified character set further except perhaps for phonetic modifier letters. Superfluous variant letters could be marked as 'to be unified' (essentially deprecated).

It would make sense to unify generally around the character subset with 'unmarked' letter names (TIFINAGH LETTER <X>, with no further precision) since, assuming it will contain YE and YO, it accurately and completely covers Northern Berber and Tuareg requirements for representing sounds. (Phonetic phenomena such as palatalisation or nasalisation are important for transcription but not a practical everyday writing system). At the moment only Moroccan IRCAM-style writing uses such 'unmarked' letters exclusively.

For writing Tuareg vowels, diacritics would be used in addition. (Which ones depend on the vowel system chosen).

Almost all Tuareg vowel systems (1)(3) can be written using the unified character set with suitable font support (with the exception of one character for Rissa Ixa's system (3), which seems to have a standalone letter (^) for short 'a' with no diacritic). The base letters for diacritics in most Tuareg vowel systems, if taken to be variant visual forms of YI/YAY, YU/YAW, YA, YEY, can be shown in the required form with a suitable font. Hawad's system (1,3,5) has an unusual side-by-side stacking mechanism of dots and macron diacritics (on base letters that look like X/Ⴟ and Σ but sound like ⵁ and Σ) that a font would need to take into account. The APT Niger system has the advantage of consistency with the IRCAM's alphabet's vowels.

I would make only one adjustment to the character set, ideally – the 'dj' sound is not universal among Tuareg variants and the letter identified as 'TIFINAGH LETTER YAJ' seems more generally to represent 'g'. The glyph contains dots, yet it is not a vowel, making it inconsistent with other letters in the set. A better code point for 'dj' would be 'TIFINAGH LETTER BERBER ACADEMY YAJ' *, to give the right default glyph with generic fonts, and this would also free YAJ up to represent Ahaggar g^y if necessary.

This adjustment would not affect IRCAM-style writing because there is no letter for 'dj' in the standard alphabet – a 'dj' sound is found in some regional variants but is written non-phonetically. One example is Saïd Belgharbi's Rifian book 'Asfiǧet'. Rather than the 'unmarked' character 'TIFINAGH LETTER YAJ' it uses the equivalent Berber Academy variant form on its cover for 'ǧ' (dj), which would be written in standard IRCAM orthography as double YAL. See (4), "Les caractéristiques phoniques de l'Alphabet Tifinage-Ircam".

A counterargument might be that web pages or documents for a particular language would have to be rendered with an appropriate regional font. However, as I have explained, to some extent this is true already (for correct display of Tuareg biconsonants, vowel dots, and leaned letters). Even if no page-embedded or locally-installed font can be identified, fallback to a generic font would show the (IRCAM) reference glyphs. Users would know most of these as they would look the same, and the rest would soon become familiar. Confusion would be unlikely in general because for each regional form the corresponding IRCAM letter is either near-identical or looks unlike any of the other regional letters, though pairs like YAH and Tuareg/Algerian YAB (ⵔ), YU and Tuareg YAW (ⵓ), YI (ⵉ) / YAY (ⵉ), and Tuareg YAD / YADD (ⴹ) are exceptions.

In Windows, Tifinagh font support used to rely on manually-installed fonts. From Windows 7, the generic fonts contain readable Tifinagh - though there are some spacing errors (already raised with Microsoft). This should give a boost to the use of the script.

To see the correct local repertoire, users would set the font used by their chat program, or set the operating system font by locale. For word processing, there would be a few commonly installed fonts per region as now (currently these are often non-Unicode). Pdf exchange allows font embedding (licences permitting). Web pages would either take advantage of common regional fonts assumed to be installed, or embed the font on the page – the technology to do this is now mature. Currently, web pages in Tuareg Tifinagh are vanishingly rare; in neo-Tifinagh they are rare and assume that the correct Unicode or non-Unicode font is installed.

The situation not covered by the unification solution would be users typing Tuareg into a web form on a page like Yahoo mail not specialized for Tuareg. The letters would appear as default IRCAM neo-Tifinagh. This is somewhat the case already, however (some features don't render correctly with generic fonts). Workarounds are to send the message as an attachment or override the browser fonts.

The most compelling counterargument would seem to be that full unification overstretches the definition of unification, because a character's variant glyphs would be substantially different from each other visually. I would argue that the variant glyphs' intended use is the same and therefore so is the fundamental identity of the corresponding character, and that the unity of the script is broken only arbitrarily by different regional letters driven by historical influences and sound changes. However, finding a good balance in Unicode is an art and each case is different.

Extension and possible further disunification

This approach would rectify current deficiencies by adding missing characters and perhaps disunifying letters which have been found to have significant variations in typography.

Each style of Tifinagh writing would encode text in a different way, so tools such as keyboards (although the layouts could be equivalent) would need to be duplicated everywhere, in areas with limited resources and in a difficult political and social environment.

The overall Tifinagh set would be large, and any further letters needed in future would have to be added and to figure in fonts' ligature substitution tables.

Conclusion

The existing support in Unicode of the Tifinagh script has some gaps. Future work could take different possible directions:-

To continue broadly as before, encoding variant letters as Unicode characters, implying that
- Characters required for writing unsupported regional styles such as Algerian neo-Tifinagh, or for writing unsupported Tuareg vowel systems, would continue to be added if needed.
- The overall Tifinagh character set would be large, yet it does not currently allow writing several Tifinagh styles with one font because of conflicting typography and lack of mandatory support for ligatures made with TIFINAGH JOINER. This conflicting typography for the same letters between styles also limits the ability to display text accurately with any arbitrary font. The problem could be overcome by disunifying.
- Each regional repertoire would need its own keyboard and other tools, all equivalent
- A large table of Tuareg ligatures would be required in fonts to cover each combination of letter variants for biconsonants.

To unify Tifinagh completely around a shared character set containing IRCAM's letters. This would imply that
- There would be no need to add new characters. This is largely true even for alternative Tuareg vowelling systems.
- Fonts would be responsible for showing the correct visual repertoire for a region, with no need to mix common and variant letters. Standard fonts would simply render the reference glyphs (which would be IRCAM's) so no change would be required to them. Specific fonts are already currently required for obtaining Tuareg effects such as filled dots, contextual shaping and ligatures.
- Typographic features like letter leaning would be handled by the fonts, not by disunifying characters. Tuareg fonts would need to map their existing typographical effects onto the new shared character set (as well as retaining them for the old variant letters for backwards compatibility).
- Only one keyboard layout would be sufficient for all regions. Local repertoires would not be constrained by the keyboard installed there, but by font availability.
- In the few situations where an alternative font to the generic font cannot be specified, a user would see the reference (IRCAM) letter forms. For example, Rifian Berber and Tamashek would appear in the same style typed on the computer command line (unless the system font is locale-dependent). Confusion is unlikely for almost all letters - see "Possible responses and implications - Full unification".

References

[1] Morocco-Canada-France (prepared by P. Andries) - Proposal to add the Tifinagh Script, n2739.pdf (Unicode proposal)

[2] Lorna A. Priest, Jon Coblentz, Andrew Savage - Proposal to encode additional Tifinagh characters , n3482.pdf (Unicode proposal)

[3] Naïma LOUALI, LES VOYELLES TOUAREGUES ET L'ALPHABET TIFINAGH : EVALUATION DE QUELQUES PROPOSITIONS RECENTES, PHOLIA (Laboratoire de Phonétique et Linguistique Africaine) Volume 8, 1993

[4] Meftaha Ameur, Abdallah Boumalk (editors) – Standardisation de l'amazighe, IRCAM, Rabat 2004

[5] Patrick Andries, Inventaire des œils de la police Hapax Berbère
http://hapax.qc.ca/polices/inventaire-des-oeils.pdf