Title:	**Prioritization of Proposals for Script Encoding**
Source:	**Rick McGowan, Michael Everson, Ken Whistler**
Date:	**November 30, 1998**
Disposition:	**Consideration by UTC**

Introduction

This document is a distillation of priorities for script encoding beyond Unicode 3.0. It attempts to answer two questions:
1. what are the script proposals that exist (or should exist)
2. what is the priority for encoding them on and off the BMP.

It is hoped that this document can provide a guide for UTC to make decisions about the relative readiness-for-publication and urgency of various proposals to be discussed after publication of Unicode 3.0. This should be revised for presentation again at the following UTC meeting, and thereafter taken up as a main piece of business following on the publication of Unicode 3.0.

The authors looked over the corpus of existing script proposals, including:
1. Proposals enumerated in the list on Michael Everson's web site,
2. Old UTC documents of "script status" from several years ago,
3. Recent "triage" documents sent to the Unicore list in October,
4. Subsequent off-line discussion,
5. Marshmallows for numerological symmetry.

The Five Categories Explained

The existing scripts and proposals have been divided into five categories, as enumerated in a subsequent section below.

Scripts in category (1) are ready to adopt as-is. Readiness for acceptance and publication decreases marginally through category five, but these categories should not be taken as a strict hierarchy of relevance. The highest priority scripts in category (2) should be given precedence in committee discussions because consensus must be reached on some pertinent questions regarding architecture or encoding model. Scripts of category (3) are as urgent as those in (2), but proposals are not as far along. Category (4) contains scripts that are well-agreed, and could be adopted with little committee involvement (they need expert scrutiny and agreement, but are not controversial as to their model or content).

It is recommended that UTC soon take up the ratification of proposals in Category (1), and following from that, discuss modelling and architectural issues pertinent to the encoding of scripts in the other categories, particu-lar-ly Hieroglypics, Math, Avestan/Pahlavi.

The Five Categories Enumerated

1. Scripts with solid proposals meeting all the criteria for immediate acceptance, unless there are real objections. Some proposals might yet need character property assignments, but otherwise are complete and stable right now, including necessary fonts for publication.

> Philippine Scripts (Tagalog, Buhid, Hanunoo, Tagbanwa; see WG2 N1933)
> Phoenician
> Buginese
> Glagolitic
> Batak

2. Scripts that are high priority to encode, the proposals for which need minor (sometimes major) work, and for which there is some high level of demand or need. These are important scholarly scripts, or are in current use.

> More Math Symbols (proposal in progress, Murray Sargent)
> Cham (proposal exists)
> Lepcha (Róng, proposal exists; scholarly contact in progress)
> Egyptian Hieroglyphs (requires model & architectural discussion)
> Tifinagh (proposal exists, needs discussion)
> Avestan (requires discussion with Pahlavi)
> Ethiopic extensions
> Old Hungarian (LTR proposal; requires discussion)

3. Scripts that are high priority to encode due to utility or demand, even if there aren't any proposals existing as yet. Those in this category now have been proposed.

> Pahlavi (no proposal; should be proposed in conjunction with Avestan)
> Newari (proposal exists; possible unification discussion)
> Lanna Tai (Tai Nua, Tai Lu & other names)

4. Lower priority scripts, with existing proposals. Some of these need very minor work, but are stable as to repertoire, overall ordering, etc. Fonts exist for many of these, and models need no discussion.

> Brahmi
> Ugaritic Cuneiform (stable)
> Old Persian Cuneiform (stable)
> Meroitic (stable)
> Shavian
> Pollard
> Deseret (stable)
> South Arabian (stable)
> Old Permic
> Linear B

5. Other scripts and proposals. This includes some minor scripts, recently-invented scripts, and controversial scripts. Some of these may be important enough to migrate to category (3) soon. At the present time, anything not in this category is not at all considered a candidate for encoding in the near future.

> Coptic (controversial disunification proposal exists)
> Blissymbols (proposal forthcoming)
> Linear A (draft proposal, expert contacted)
> Sinaitic
> Tengwar (a Tolkein script; proposal exists)
> Cirth (a Tolkein script; proposal exists)
> Phaistos disk (undeciphered)
> Klingon (a Fictional script; stable; have a letter from the KLI)
> Indus Valley (no proposal, but well defined and expert contacted)
> Pahawh Hmong (no proposal, but expert contacted)
> Aramaic (no proposal)
> Samaritan (no proposal)
> Mandaean (no proposal but expert contacted)
> 'Phags-pa (no proposal)
> Siddham (no proposal)
> Javanese (one older partial proposal exists)
> Kayah Li (no proposal)
> Sumero-Akkadian Cuneiform (no proposal)
> Mayan Hieroglyphs (no proposal)
> Unifon (no proposal)
> Cypriot
> Aegean Numbers