

Encoding SignWriting in the Universal Character Set

by Michael Everson
Dublin, Ireland, August 15, 1999



The Universal Character Set or UCS (International Standard ISO/IEC 10646-1 and computer industry implementation Unicode) is a solution to the many problems which earlier computer coding systems have had when dealing with multilingual text. 8-bit computer systems such as Mac OS 8 and Windows 95 had to resort to complex solutions for switching between character sets (usually understood by the user as "code pages" or "fonts") even to do such simple things as display German and Polish text in the same document. The UCS presents a single coding system for text that can handle, to put it simplistically, all the letters of all the alphabets of all the languages of the world. Use of this standard is being given the highest priority by such companies as Microsoft, Apple, and Sun. Today, the UCS handles writing systems as different as Latin, Greek, Cyrillic, Armenian, Chinese, Bengali, Tibetan, Arabic, Hebrew, Korean, Mongolian, and Cherokee. Plans to extend the repertoire to include scripts like Egyptian Hieroglyphs, Phoenician, Aramaic, Tagalog, and Blissymbols are being considered by the Technical Committees responsible for ISO/IEC 10646 and the Unicode Standard.

SignWriting, as a new writing system uniquely suited to the representation of the world's Sign Languages, has recently come to the attention of these Technical Committees. Michael Everson, Irish national representative to the committee responsible for ISO/IEC 10646, is a director of an Irish company dedicated to providing computing solutions for lesser-used languages. He has spearheaded discussions with the Deaf Action Committee to explore the possibility of encoding SignWriting in the UCS.

The advantage of having SignWriting encoded in the UCS are many. As the corpus of Sign Language literature, currently written using the DAC's SignWriter 4.0 program, grows, it becomes increasingly important that users have the ability to exchange data safely (with no corruption) and to search, sort, and otherwise manipulate this data in the same way that users of spoken languages do. With the increasing availability of SignWriting text on the internet, this is becoming even more relevant. Since it is clear that the Universal Character Set will be the means for encoding and exchanging text, supported by all software vendors, it is important that work begin now to prepare SignWriting for encoding in the UCS.

SignWriting has a number of features in common with other writing systems encoded in the UCS. It also has unique features, such as the semantically-relevant reversal (flopping) and rotation of base characters, and their positioning in both vertical and horizontal contexts, which will require a concerted effort on the part of programmers involved with the SignWriter program, experts in UCS implementation, and the user community. Preliminary explorations of the effort required indicate that 6-10 man-years of effort will be required. Funding for this effort is urgently required if SignWriting is to take its place among the world's writing systems in the UCS. This will benefit all users of SignWriting in the countries currently using SignWriting and will in addition pave the way for expansion of SignWriting into new communities, as it is well-established that SignWriting is a robust writing system suitable for representing Sign Languages worldwide.



SignWriting is Helping Bring Recognition to the World's Sign Languages

- There are millions of people worldwide who communicate daily in a Sign Language.
- SignWriting is a way to read, write and type all the different Sign Languages in the world.
- The International Organization for Standardization (ISO) sets the world's standards for the representation of languages in computers.
- Computer programmers refer to the “ISO 639-2 Registration Authority” for the standard codes used to represent the world's languages. This helps standardize software development. Before February, 2000, Sign Languages were not included in this listing.
- In September, 1999, the Deaf Action Committee for SignWriting (the DAC), and the Irish National Body applied to the Registration Authority, requesting that the world's Sign Languages be included. The application was approved in February, 2000!
- The Universal Character Set (UCS) is an international standard for representing languages in computers known as ISO/IEC10646, or **Unicode**. It will take several years to prepare SignWriting for encoding in the Universal Character Set.
- The project: **Encoding SignWriting in the Universal Character Set** needs funding.
- Encoding SignWriting in the UCS will make it easier for software developers to incorporate SignWriting into their programs. It will facilitate the sharing of Sign Language data between computers and across the internet.
- The result? **Deaf people worldwide will have improved and increased communication!**