

Accredited Standards Committee¹

NCITS, Information Processing Systems

Doc: **L2/97-107**
Date: 27 May 1997
Author: John H. Jenkins
Apple Computer, Inc.
2 Infinite Loop
M/S 302-2IS
Cupertino CA 95014
Phone: (408) 974-6276
Fax: (408) 862-4566
e-mail: jenkins@apple.com

Proposal to add the Linear B to Unicode/ISO-IEC 10646

The script called *Linear B* is a syllabic system that was used on the island of Crete (and parts of the nearby mainland) to write the oldest recorded variety of the Greek language. Linear B clay tablets predate Homeric Greek by some 700 years, the latest being from about 1375 BC. Major archaeological sites include Knossos, first uncovered in about 1900 by Sir Arthur Evans, and a major site near Pylos on the mainland. The majority of inscriptions currently known are inventories of commodities and accounting records.

The script resisted early attempts at decipherment, but it finally yielded to the efforts of Michael Ventris, an architect and amateur decipherer. Ventris' breakthrough in decipherment came after the realization that the language might be Greek, and not (as had been previously thought) a completely unknown language. Ventris formed an alliance with John Chadwick, and decipherment proceeded quickly. Ventris and Chadwick published a joint paper in 1953.

Linear B was written from left to right with no non-spacing marks or other complications. The script consists mainly of a number of phonetic signs representing the combination of a consonant and vowel. There are 60 known phonetic signs, a few signs that seem to be mainly free variants (Chadwick's *optional signs*), a few unidentified signs, numerals, and a number of ideographic signs which were used mainly as counters for commodities. Some ligatures formed from combinations of syllables were apparently used as well. Chadwick gives several examples of these ligatures, which are not included in this encoding.

The signs having phonetic values beginning with *J* are pronounced in the German manner as the English *Y*.

The current road map for WG2 recommends encoding Linear B in Plane 1. This proposal reflects that.

¹ Operating under the procedures of the American National Standards Institute
NCITS Secretariat, Information Technology Industry Council
1250 Eye Street NW, Suite 200, Washington DC 20005-3922 Telephone: 202-737-8888 (press 1 twice) Fax: 202-638-4922

Differences from TR3

The Linear B proposal here differs from the one authored by Rick McGowan and issued as part of *Technical Report #3* by the Unicode Consortium.

- The block of unidentified symbols has been moved so that it doesn't intervene between *two* and *twe*. The order of the symbols in column 0001 $\times x4y$ has been modified somewhat.
- The old syllables *pa3*, *swa*, and *swi* are marked as optional with a question mark in Chadwick (1987) and considered unidentified by Hooker (1980). They've been moved to the unidentified section accordingly.
- The unidentified symbols (13 in count) are in a column by themselves. Their name has been changed to LINEAR B SYLLABLE NN, where NN is the standard Linear B syllable number.
- The names have been changed to remove the parenthetical identifications.
- The symbols for numerals have been added, and the mensuration symbols separated from the other ideograms are placed immediately after the numerals. There are no other ideograms included in this proposal.

Issues

Unification with Linear A is still moot, although the evidence would tend to be against it.

The case for unification is that the two sets have considerable overlap, and when the Linear A syllabary is read using the Linear B sounds, place and personal names are identifiable.

Chadwick (1987) uses the analogy of Cyrillic and Latin to urge caution, however. If Cyrillic were deciphered and Latin not, a number of proper nouns written in the Latin alphabet could still be seen to make sense if the Cyrillic sounds were used—but the two alphabets are very different nonetheless.

It would seem best to frame the proposal so that unification is neither forced nor precluded.

(The numeral systems, however, of both scripts are completely known and there is no question but that they should be unified.)

Information on the ligatures and ideograms is still being gathered. It was felt best to proceed with encoding the syllabary, however, while this work is being done.

Linear B numerals consist of a set of five symbols which are repeated as needed to indicate the proper number. It would be like writing “11111” for “5.” The full set of Linear B numerals could be encoded using the five symbols indicated here (for units, tens, hundreds, thousands, and myriads). There is some sentiment, however, for encoding each possible digit separately (1, 2, 3, ..., 9, 10, 20, 30, ..., 90, ..., 90000), which would require 45 code points.

Linear B uses a short vertical stroke to separate words. There is a possibility that other word separation characters might be required.

Some Sources

Bennett, Emmett L. (1996). "Aegean Scripts." In *The World's Writing Systems*, edited by Peter

T. Daniels and William Bright. Oxford: University Press.

Chadwick, John. (1970). *The Decipherment of Linear B*. Second Edition. Cambridge: University Press.

Chadwick, John. (1987). *Linear B and Related Scripts*. Berkeley: University of California Press.

Hooker, J.T. (1980). *Linear B: An Introduction*. Bristol, Bristol Classics Press.

Sampson, Geoffrey. *Writing Systems; a linguistic introduction*.

Ventris, Michael, and Chadwick, John. (1959). *Documents in Mycenaean Greek*. Cambridge: University Press.

Structure of the current proposal

Basic syllabary

Positions 0001 `xx00` through 0001 `xx41` are used for the 13 x 5 basic matrix of Mycenaean syllables written with Linear B. Note that not all the positions are used. There is basic agreement on these characters, with the exception of 0001 `xx0E` LINEAR B SYLLABLE JU, which is omitted by Hooker for some reason.

Optional syllabary

Positions 0001 `xx42` through 0001 `xx4F` are used for a set of additional syllables which are reasonably well-understood and were used to provide alternate spellings or occasional spelling of sounds not handled by the basic set. There is less general agreement on this set. In particular, we omit three syllables which Chadwick (1987) identifies with a question mark.

Unidentified syllables

Positions 0001 `xx50` through 0001 `xx5F` are used for elements of the Linear B syllabary which are as yet unidentified.

Numerals

Positions 0001 `xx60` through 0001 `xx6F` are used for Linear B numerals. There is no disagreement on how numerals were written with Linear B.

Note that this set is a subset of that used for Linear A. Inasmuch as the code points 0001 `xx60` through 0001 `xx64` are the only ones used as the moment, the additional Linear A numerals could fit in the remainder of the column.

Punctuation

Positions 0001 `xx70` is used for the word separator. A gap of six code points is left between the word separator and the mensuration characters to provide room for other punctuation or ligation

symbols.

Measures

Positions 0001 xx77 through 0001 xx7F are used for mensuration symbols. There is no disagreement on the symbols used for mensuration or their general ordering. There is some minor disagreement on the exact values and the exact interrelationship between the various subunits. Linear B includes symbols for measuring weight, dry volumes, and liquid volumes. Two symbols are used for both liquid and dry volumes.

Ideograms

Linear B used a large number of ideograms to indicate specific objects: people, animals, containers, and so on. The current proposal includes no ideograms. It is assumed that the ideograms will be encoded beginning with code point 0001 xx80 and run (perhaps) for a couple of hundred code points.

Just to give a general overview of the issues:

Bennet (1996) shows some three dozen ideograms.

Chadwick (1987) shows about two dozen.

Hooker (1980) lists the sixty most common.

Chadwick (1959) lists roughly one hundred fifty.

A complete list of the ideograms has, in fact, been provided the Aegean Scripts ad hoc committee, but it didn't arrive in time to fully integrate it into this proposal.

Names

```
0001 xx00 LINEAR B SYLLABLE A
0001 xx01 LINEAR B SYLLABLE E
0001 xx02 LINEAR B SYLLABLE I
0001 xx03 LINEAR B SYLLABLE O
0001 xx04 LINEAR B SYLLABLE U
0001 xx05 LINEAR B SYLLABLE DA
0001 xx06 LINEAR B SYLLABLE DE
0001 xx07 LINEAR B SYLLABLE DI
0001 xx08 LINEAR B SYLLABLE DO
0001 xx09 LINEAR B SYLLABLE DU
0001 xx0A LINEAR B SYLLABLE JA
0001 xx0B LINEAR B SYLLABLE JE
0001 xx0C (This position shall not be used)
0001 xx0D LINEAR B SYLLABLE JO
0001 xx0E LINEAR B SYLLABLE JU
0001 xx0F LINEAR B SYLLABLE KA

0001 xx10 LINEAR B SYLLABLE KE
0001 xx11 LINEAR B SYLLABLE KI
```

0001 xx12 LINEAR B SYLLABLE KO
0001 xx13 LINEAR B SYLLABLE KU
0001 xx14 LINEAR B SYLLABLE MA
0001 xx15 LINEAR B SYLLABLE ME
0001 xx16 LINEAR B SYLLABLE MI
0001 xx17 LINEAR B SYLLABLE MO
0001 xx18 LINEAR B SYLLABLE MU
0001 xx19 LINEAR B SYLLABLE NA
0001 xx1A LINEAR B SYLLABLE NE
0001 xx1B LINEAR B SYLLABLE NI
0001 xx1C LINEAR B SYLLABLE NO
0001 xx1D LINEAR B SYLLABLE NU
0001 xx1E LINEAR B SYLLABLE PA
0001 xx1F LINEAR B SYLLABLE PE

0001 xx20 LINEAR B SYLLABLE PI
0001 xx21 LINEAR B SYLLABLE PO
0001 xx22 LINEAR B SYLLABLE PU
0001 xx23 LINEAR B SYLLABLE QA
0001 xx24 LINEAR B SYLLABLE QE
0001 xx25 LINEAR B SYLLABLE QI
0001 xx26 LINEAR B SYLLABLE QO
0001 xx27 (This position shall not be used)
0001 xx28 LINEAR B SYLLABLE RA
0001 xx29 LINEAR B SYLLABLE RE
0001 xx2A LINEAR B SYLLABLE RI
0001 xx2B LINEAR B SYLLABLE RO
0001 xx2C LINEAR B SYLLABLE RU
0001 xx2D LINEAR B SYLLABLE SA
0001 xx2E LINEAR B SYLLABLE SE
0001 xx2F LINEAR B SYLLABLE SI

0001 xx30 LINEAR B SYLLABLE SO
0001 xx31 LINEAR B SYLLABLE SU
0001 xx32 LINEAR B SYLLABLE TA
0001 xx33 LINEAR B SYLLABLE TE
0001 xx34 LINEAR B SYLLABLE TI
0001 xx35 LINEAR B SYLLABLE TO
0001 xx36 LINEAR B SYLLABLE TU
0001 xx37 LINEAR B SYLLABLE WA
0001 xx38 LINEAR B SYLLABLE WE
0001 xx39 LINEAR B SYLLABLE WI
0001 xx3A LINEAR B SYLLABLE WO
0001 xx3B (This position shall not be used)
0001 xx3C LINEAR B SYLLABLE ZA
0001 xx3D LINEAR B SYLLABLE ZE
0001 xx3E (This position shall not be used)
0001 xx3F LINEAR B SYLLABLE ZO

0001 xx40 (This position shall not be used)
0001 xx41 LINEAR B SYLLABLE HA
0001 xx42 LINEAR B SYLLABLE INITIAL AI
0001 xx43 LINEAR B SYLLABLE INITIAL AU

0001 xx44 LINEAR B SYLLABLE DWE
0001 xx45 LINEAR B SYLLABLE DWO
0001 xx46 LINEAR B SYLLABLE NWA
0001 xx47 LINEAR B SYLLABLE PHU
0001 xx48 LINEAR B SYLLABLE PTE
0001 xx49 LINEAR B SYLLABLE RJA
0001 xx4A LINEAR B SYLLABLE RAI
0001 xx4B LINEAR B SYLLABLE RJO
0001 xx4C LINEAR B SYLLABLE TA2
0001 xx4D LINEAR B SYLLABLE TWE
0001 xx4E LINEAR B SYLLABLE TWO
0001 xx4F (This position shall not be used)

0001 xx50 LINEAR B SYMBOL 18
0001 xx51 LINEAR B SYMBOL 19
0001 xx52 LINEAR B SYMBOL 22
0001 xx53 LINEAR B SYMBOL 34
0001 xx54 LINEAR B SYMBOL 47
0001 xx55 LINEAR B SYMBOL 49
0001 xx56 LINEAR B SYMBOL 56
0001 xx57 LINEAR B SYMBOL 63
0001 xx58 LINEAR B SYMBOL 64
0001 xx59 LINEAR B SYMBOL 79
0001 xx5A LINEAR B SYMBOL 82
0001 xx5B LINEAR B SYMBOL 83
0001 xx5C LINEAR B SYMBOL 86
0001 xx5D (This position shall not be used)
0001 xx5E (This position shall not be used)
0001 xx5F (This position shall not be used)

0001 xx60 LINEAR B NUMERAL UNITS
0001 xx61 LINEAR B NUMERAL TENS
0001 xx62 LINEAR B NUMERAL HUNDREDS
0001 xx63 LINEAR B NUMERAL THOUSANDS
0001 xx64 LINEAR B NUMERAL MYRIADS
0001 xx65 (This position shall not be used)
0001 xx66 (This position shall not be used)
0001 xx67 (This position shall not be used)
0001 xx68 (This position shall not be used)
0001 xx69 (This position shall not be used)
0001 xx6A (This position shall not be used)
0001 xx6B (This position shall not be used)
0001 xx6C (This position shall not be used)
0001 xx6D (This position shall not be used)
0001 xx6E (This position shall not be used)
0001 xx6F (This position shall not be used)

0001 xx70 LINEAR B WORD SEPARATOR
0001 xx71 (This position shall not be used)
0001 xx72 (This position shall not be used)
0001 xx73 (This position shall not be used)
0001 xx74 (This position shall not be used)
0001 xx75 (This position shall not be used)

0001 xx76 (This position shall not be used)
 0001 xx77 LINEAR B WEIGHT BASE UNIT
 0001 xx78 LINEAR B WEIGHT FIRST SUBUNIT
 0001 xx79 LINEAR B WEIGHT SECOND SUBUNIT
 0001 xx7A LINEAR B WEIGHT THIRD SUBUNIT
 0001 xx7B LINEAR B WEIGHT FOURTH SUBUNIT
 0001 xx7C LINEAR B DRY MEASURE FIRST SUBUNIT
 0001 xx7D LINEAR B LIQUID MEASURE FIRST SUBUNIT
 0001 xx7E LINEAR B MEASURE SECOND SUBUNIT
 0001 xx7F LINEAR B MEASURE THIRD SUBUNIT

Code Chart

	0	1	2	3	4	5	6	7
0								
1								
2								
3								
4								
5								
6								
7								
8								
9								
A								
B								
C								
D								
E								
F								