Title: Proposal to encode the Northern Palaeohispanic script

From: Joan Ferrer i Jané, Grup LITTERA, Universitat de Barcelona; Noemí Moncunill, Grup LITTERA, Universitat de Barcelona; Javier Velaza, Grup LITTERA, Universitat de Barcelona and Deborah Anderson, SEI, UC Berkeley

Date: 1st January 2020

This is a revised proposal for Northern Palaeohispanic. Changes from the earlier proposal L2/19-332 are identified by yellow highlighting in this proposal. Changes include one new character, two new figures, a new bibliographic reference, improved captions, and some corrections.

Earlier documents include:

L2/19-332 Proposal to encode the Northern Palaeohispanic script

L2/19-045 New charts for Northern and Southern Palaeohispanic

The following documents were based on a unified script of Northern and Southern Palaeohispanic.

L2/18-283 Proposal to encode the Palaeohispanic script

L2/18-030 Proposal to encode the Palaeohispanic script

L2/17-129 Proposal to encode the Palaeohispanic script

The two preliminary proposals below were only for what is later called "Northern Palaeohispanic":

L2/15-120 Preliminary proposal to encode the north-eastern Iberian script

L2/15-012 Preliminary proposal to encode the north-eastern Iberian script

From: Joan Ferrer i Jané, Grup LITTERA, Universitat de Barcelona; Noemí Moncunill, Grup LITTERA, Universitat de Barcelona; Javier Velaza, Grup LITTERA, Universitat de Barcelona and Deborah Anderson, SEI, UC Berkeley

Date: 1st January 2020

Title: Proposal to encode the Northern Palaeohispanic script

1. Summary

Palaeohispanic scripts are attested in the Iberian Peninsula by *ca*. 2,700 inscriptions dating from the 7th century BC to the 1st century AD. They were used to write at least four different local languages: Celtiberian, Iberian, the south-western or Tartessian language and probably as well Turdetanian. Although the longest inscription contains *ca*. 500 words, most long texts rarely reach over 50 words; others, on the other hand, are very short and contain just a personal name or abbreviations.

The Palaeohispanic script family consists of several scripts that can be divided into 2 types: the northern (with *ca*. 2,500 inscriptions) and the southern group (with *ca*. 170 inscriptions). That's why we codify them according to two different standards:

1. Northern Palaeohispanic, which includes the northeastern Iberian and the Celtiberian script.

2. Southern Palaeohispanic, which includes the south-eastern Iberian script, the south-western or Tartessian script, the Turdetan script and the Espanca abecedary.

All Palaeohispanic writing systems are characterised by a similar corpus of signs and by the coexistence of alphabetic and syllabic characters. Moreover, all of them share a common ancestor, which might ultimately arise from the Phoenician alphabet (see fig. 1). However, the differences between the two groups are too deep to be appropriately processed into a unique Unicode character set. The main obstacle is the different degree of decipherment between the two groups, which is almost complete for the northeastern Iberian group and still incomplete for at least a third of the signs attested in the southern scripts. The second obstacle is that a large number of signs shared by these two groups actually have different values: it is the case for most of the vocalic signs, as well as for other frequent signs for which there is clear consensus on their value.

The decipherment of the northern Palaeohispanic scripts was accomplished at the beginning of the 20th century by Manuel Gómez-Moreno (1922, 1949). Nevertheless, some aspects were not entirely deciphered until very recent dates. Such is the case of a variant of this script called *the dual system*, which consists of the use of signs with two variants, each of them with its own distinctive value, differing from each other in presenting an additional stroke (e. g. Χ = **da** and Χ̄ = **ta**). In fact, some concrete features of this system are even still undergoing research. Unlike the northern scripts, the southern scripts have not been fully deciphered, since there are many signs for which there is no agreed value among specialists. The non-deciphered characters will be named in this proposal after an arbitrary code, as it is currently in use in the specialized bibliography.

The Palaeohispanic inscriptions are being edited and digitalized in the Hesperia open access database (http://hesperia.ucm.es/), in the framework of a research project carried out by a team of scholars from different Spanish universities.

2. Background

This proposal, together with the one devoted to "Southern Palaeohispanic", is an updated version of three earlier proposals: "L2/15-120-Preliminary proposal to encode the northeastern Iberian script";

"L2/15-119 Preliminary proposal to encode the southern Palaeohispanic scripts"; and L2/18-283 "Proposal to encode the Palaeohispanic script".

The last of them was an attempt to codify all the Palaeohispanic scripts together, under a single encoding. However, after submitting this possibility for analysis with several Unicode specialists, we have come to the conclusion that a separate codification for the southern and northern scripts makes it possible to better reproduce the graphematic system in the different scripts. The revised charts in the current proposal now reflect repertoires that reflect the graphemes in the writing systems. Therefore, two documents are produced, this one for the northern Palaeohispanic script, and another one for the southern Palaeohispanic script.

With the proposed repertoires, all characters in Northeastern Iberian and Celtiberian scripts can be represented. As with Old Italic, fonts will be used to represent the different alphabets of Iberia.

The earlier proposals, L2/15-120 and L2/15-119, contain detailed discussion on the characters, figures, and references. See also the article J. Ferrer, N. Moncunill, J. Velaza, "Toward a systematisation of Palaeohispanic scripts in Unicode: synthesising multiple transcription hypotheses into two consensus encodings", *Palaeohispanica* 15, 2015, pp. 13-55.

3. Structure

All Palaeohispanic scripts are semisyllabic: vowels, nasals, laterals, sibilants and trills are alphabetical, whereas characters for plosives are syllabic.

There is another feature shared by most Palaeohispanic scripts: the possibility they offer to differentiate some similar signs with close phonetic value by an additional stroke; rather than a mere diacritic (although originally it might have been so), this stroke tends to be an integrating component of the sign itself. This subset of scripts with a larger number of variants has been labelled as "dual". The recent discovery of dual abecedaries confirms that these dualities were integrated in the standard scripts, where the pair of signs appears always in the same order: the complex variant, with its additional stroke, in the first place and, thereafter, the simple one.

These dualities can affect different sets of sounds: plosives, continuous consonants and even vowels. However, and despite their autonomous apparition in the abecedaries, only for the first ones the phonetic opposition (in this case, voiced – voiceless) between the simple and the complex has been confirmed. Therefore, according to the extension of the use of dualities, it is possible to identify different subsets within every script:
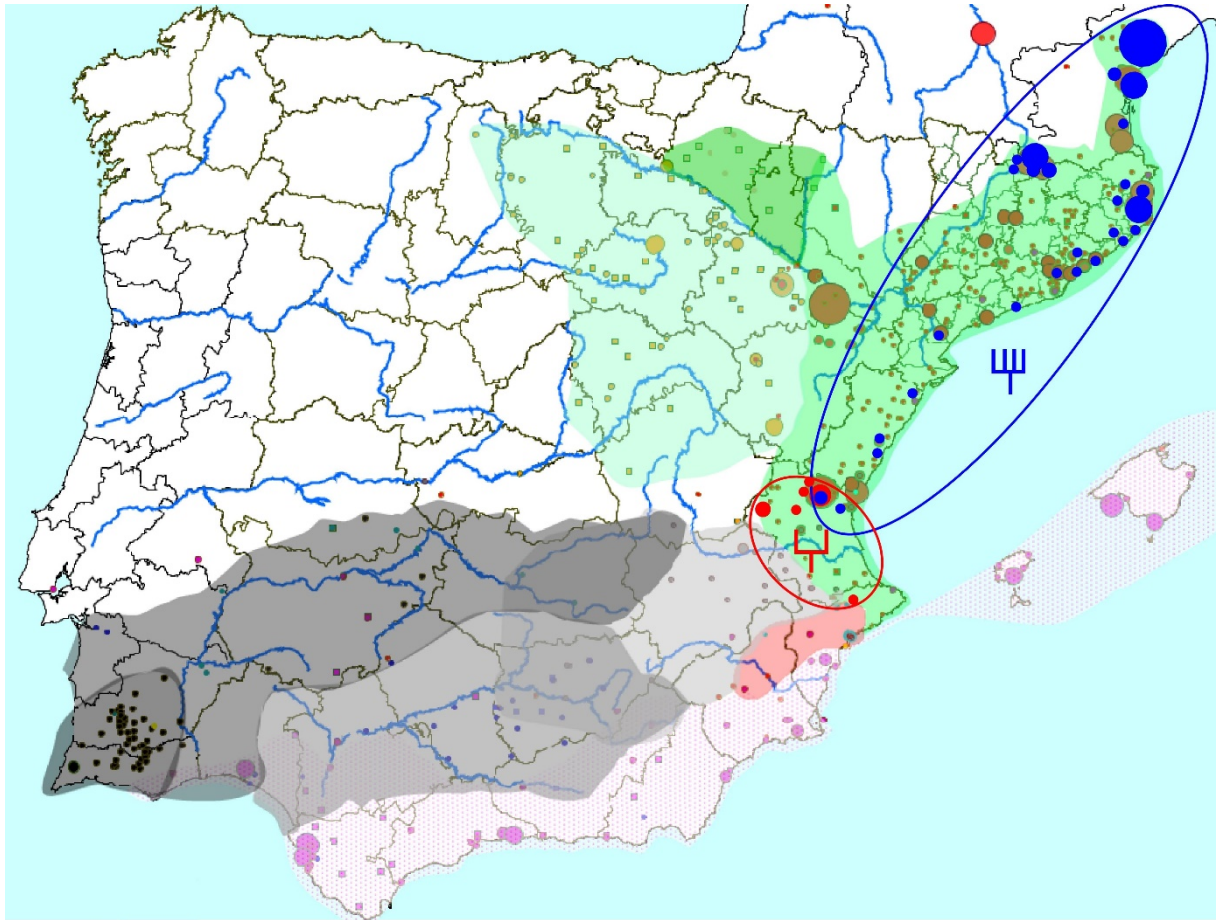
- The northeastern Iberian script can be divided into: i. extended dual, when dualities affect plosives, continuous consonants and vowels; ii. standard dual, when it only affects plosives, and iii. non-dual, when the script lacks dualities.
- The Celtiberian script, either in its eastern or western variant, can be divided into: i. standard dual, or ii. non-dual.

The standard script for Unicode has been built taking into account an inventory of signs as large as possible, including all dual variants confirmed in the extant abecedaries; in general terms, the glyphs for the proposed characters to be encoded match the glyphs of the northeastern Iberian dual extended script.

In addition, the Unicode repertoire also considers as meaningful a three-elements variability for the signs **ke** and **ka** in the standard or extended dual script: ₵/₵/₵ and Ѧ/Ѧ/Ѧ. Indeed, some long standard dual inscriptions show the simultaneous use of three variants of the **ke** sign. That's the case of the lead sheet from Castelló (F.6.1) where two-stroke **ke** (₵) coexist with one-stroke (₵) and no-marked variants

(ᒉ). In a similar way, on the lead sheet from Ensérune (B.1.373*; Hesperia HER.02.773) a two-dot variant of the **ke** sign (ᒉ), coexists with a one-dot (ᒉ) and a no-marked variant (ᒉ). Furthermore, it needs to be pointed out that the most common word in the lead sheet from Castelló with the two-stroke **ke**, **balke**, was written in this same way in painted inscriptions of the extended dual type from Llíria. This behavior can also be observed in the lead sheet from Los Villares (F.17.2) where three simultaneous variants of the **ka** sign are used (Λ/Λ/Λ). This suggests that the three-element variation for these two characters is also a characteristic of the extended dual script, although it is not explicitly documented in any of the known abecedaries.

The number of significant variants is a critical issue in the definition of each writing; from our point of view, the proposed codification cannot leave aside epigraphic evidence, based on the point of view of the original user, and add characters that did not exist as autonomous forms. Although a superficial look at the corpus may suggest some of these characters might have existed, this is caused by the fact that we are unifying under a single codification different writings and, within them, different epigraphic schools. The main variants used by the different scripts and epigraphic schools will be expressed by means of different fonts which will allow the variability of writing to be better represented.

That's the case of the **ti** (Ψ / Ψ / Ψ) and **to** (Ш / Ш / Ш) characters, which show an apparently three-elements variation; however, the extant abecedaries, which represent the real set of signs, confirm that there are only two meaningful signs, that is only a two-elements distinction[1]. The most common dual opposition found in Palaeohispanic inscriptions, used almost in all the Iberian territory (blue dots in the map of fig. 0), is expressed as Ψ / Ψ and Ш / Ш, as confirmed in the Tos Pelat abecedary (Fig. 5). There is, however, a residual use where the opposition is expressed as Ψ / Ψ and Ш / Ш, as shown in the Castellet de Bernabé's abecedary (Fig. 4). This exceptional use is mainly found in a small area around Llíria (València, red dots in the map). In this tradition, the glyphs Ψ/Ш express the value represented by Ψ/Ш in the most common tradition and the glyphs Ψ/Ш express the value represented by Ψ/Ш in the most widespread tradition. That's why our proposal only encodes two values corresponding to the main glyphs used in the most common tradition (Ψ / Ψ and Ш / Ш). Regarding the above mentioned residual tradition in Llíria and surroundings, the glyphs Ψ / Ш will be always codified as Ψ and Ш, as they always represent the simple value. On the other hand, the glyphs Ψ / Ш will be codified always as they look like, except when they coexist in the same inscription with Ψ/Ш (note however that this is a very rare situation, which occurs up till now only once in the corpus). In this particular case, in order not to lose the graphematic opposition between the marked and unmarked characters (expressed here as Ψ / Ψ and Ш/Ш, instead of the standard Ψ / Ψ and Ш / Ш), the glyphs Ψ / Ш will be codified as Ψ / Ш, as shown in fig. 3.

A similar thing occurs with other characters which apparently seem to display a more than two elements variation, such as the vowel **o** Ν / Η / Η / Η. In this case, the signs with more than one additional stroke have not been encoded separately, since they can be always considered an allograph of the glyph with only one additional bar. The opposition between the marked and unmarked variant of this character is scarcely documented: three times in abecedaries and only twice in other types of inscriptions. The Tos Pelat's abecedary contains at least two abecedaries (Fig. 5), partially in the form of a palimpsest, all of them displaying only two meaningful signs of the vowel **o**, which is consistent with what happens with the rest of vowels (see Fig. 5). However, one of the Tos Pelat's abecedaries

---

[1] Ferrer i Jané 2019.

Fig. 0.- Geographic distribution of the less common variants of the **ti** sign. Blue dots (four strokes variants). Red dots (two stroke variants).

4. Direction of script

The proposed default direction of the script is left to right, which is the predominant direction in inscriptions of the northern group. In order to render texts right to left, users should use RLO and PDF overrides, or other mechanisms as described in UAX #9 (http://www.unicode.org/reports/tr9/proposed.html). If the default direction of the script is overridden, the glyphs in the font should be mirrored from those presented here.

---

[2] Ferrer i Jané 2015, fig. 3 and 6.

5. Character names

The character names are based by default on the northeastern script, which is the script with more inscriptions and the one whose decipherment is completed. However, since the values are not usually the same in each Palaeohispanic script, notes are used to specify the value of the character in every script. For instance:

The lack of notes implies that the sign has the same value in all scripts. For instance:

1022B ⋈ PALAEOHISPANIC NORTHERN LETTER N1

When a sign is exclusive of one single script it is specified that way:

10234 ∟ PALAEOHISPANIC NORTHERN NUMERAL A

- Northern Palaeohispanic

The marked signs in the "dual" scripts are named with the terminology "with additional stroke" or " with two additional strokes ". For instance:

10202 ⱶ PALAEOHISPANIC NORTHERN LETTER E WITH ADDITIONAL STROKE

- Northeastern Iberian (Dual Extended)

1020F ⋀ PALAEOHISPANIC LETTER KA WITH TWO ADDITIONAL STROKES

- Northeastern Iberian (Dual)

When the phonetic value of a sign is unknown, we use the conventional code used in the specialized bibliography. For instance:

10232 ⩔ PALAEOHISPANIC NORTHERN LETTER S87

Please note that the names of the letters don't always reflect the transcription system used in the discursive parts of the proposal. These are the correspondences:

S1 = **s**
S2 = **ś**
S1 WITH ADDITIONAL STROKE = **ŝ**
N1 = **n**
N1 WITH ADDITIONAL STROKE = **ń**
N2 = **m**
N3 = **ḿ**
N4 = **m̌**
R1 = **r**
R2 WITH ADDITIONAL STROKE = **ř**
R2 = **ŕ**
A = **a**

A WITH ADDITIONAL STROKE = **á**
A2 = **â**
E = **e**
E WITH ADDITIONAL STROKE = **é**
I = **i**
I WITH ADDITIONAL STROKE = **í**
O = **o**
O WITH ADDITIONAL STROKE = **ó**
U = **u**
U WITH ADDITIONAL STROKE = **ú**
KA = **ga**
KA WITH ADDITIONAL STROKE = **ka**
KA WITH TWO ADDITIONAL STROKES = **ḱa**
KE = **ge**
KE WITH ADDITIONAL STROKE = **ke**
KE WITH TWO ADDITIONAL STROKES = **ḱe**
KI = **gi**
KI WITH ADDITIONAL STROKE = **ki**
KO = **go**
KO WITH ADDITIONAL STROKE = **ko**
KU = **gu**
KU WITH ADDITIONAL STROKE = **ku**

…

6. Numbers

Iberian metrological expressions are basically formed by groups of vertical bars (equivalent to the sign **ba**) to generate the numerical component of the expression: I = 1, II = 2, III = 3, IIII = 4, IIIII =5. The accumulation of bars can reach up to 20 elements (F.17.1). Occasionally these bars can appear together with a sign similar to Greek **Π**, which appears to be acting as an auxiliary base, perhaps with the value of 5 or 6.

Some metrological expressions use a specific **L**-shaped sign, which does not match any other character of the Iberian script; the numerical value for that sign is still uncertain. This sign also appears in metrological expressions on painted amphora inscriptions from Vieille-Toulouse (for instance **L III**) and in lead-sheet inscriptions from Ìatova (for instance **L Π IIIII** [F.20.2]).

Finally, some coin inscriptions present value marks, which, in some cases, have an equivalent symbol formed by the initial of the unit followed by the numerical component. In the case of **undikesken** coins, quarters show the – sign and halves the **=** sign, which is actually a reduplication of the former (¼ + ¼ = ½).

7. Punctuation

The most common word separator consists of two vertical dots. Nevertheless, the oldest epigraphic tradition tends to use rather 3 or more vertical dots; in the most recent inscriptions on stone, on the other hand, the use of an isolated dot is frequent, imitating the Roman style. Finally, the vertical bar can also be used and, in some rare cases, just a blank.

8. Order

For the code chart: vowels will appear in the alphabetical order **a**, **e**, **i**, **o**, **u**; plosives in the usual alphabetical order **b**, **k/g**, **t/d**; and continuous consonants in the alphabetical order **l**, **m**, **n**, **r**, **s**. The marked-sign pairs will be grouped together, the marked character preceding the unmarked, as appears in the northeastern Iberian abecedaries. The conflictive **T**-shaped sign is grouped together with nasals, as it actually appears in the attested abecedaries. The conflictive sign in the shape of an Iberian l (â) with an additional stroke is placed together with this sign, as it appears in the Castellet de Bernabé's abecedary. Numerals are grouped at the end after letters.

The proposed order for sorting is as follows: **a**, á, **â, ba**, **be**, **bi**, **bo**, **bu**, **da, ta, de, te, di, ti, do, to, du, tu**, **e**, **é, ga, ka, Ḱa, ge, ke, Ḱe, gi, ki, go, ko, gu, ku**, **i**, í, **l**, **m, n**, **o**, ó, **r**, ŕ, ř, ś, **s**, ŝ, **u**, ú, ḿ, **m̌, S87**. Specific exceptions to the alphabetical order are as follows:

- Consecutive order for simple sibilant (**s**) and sibilant with an additional stroke **(ŝ)**;

- Consecutive order for voiceless and voiced plosives in order to keep together the dual and non-dual transcriptions of the same elements (for instance, the word **ekiar** / *egiar*).

- Consecutive order for **m** and **n**, since they are signs that can alternate (for instance **iunstir** / **iumstir**).

- Consecutive order for the supposed nasal **ḿ** and **m̌**, after the two signs for **u**, since the characteristic vocalic component of **ḿ** can be usually identified as **u** (for instance **ḿbaŕ** / VMAR).

9. Unicode Character Properties

All the Palaeohispanic letters from 10200 to 10238 are as below.

10200;NORTHERN PALAEOHISPANIC LETTER A WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10201;NORTHERN PALAEOHISPANIC LETTER A;Lo;0;L;;;;;N;;;;;

10202;NORTHERN PALAEOHISPANIC LETTER E WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10203;NORTHERN PALAEOHISPANIC LETTER E;Lo;0;L;;;;;N;;;;;

10204;NORTHERN PALAEOHISPANIC LETTER I WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10205;NORTHERN PALAEOHISPANIC LETTER I;Lo;0;L;;;;;N;;;;;

10206;NORTHERN PALAEOHISPANIC LETTER O WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10207;NORTHERN PALAEOHISPANIC LETTER O;Lo;0;L;;;;;N;;;;;

10208;NORTHERN PALAEOHISPANIC LETTER U WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10209;NORTHERN PALAEOHISPANIC LETTER U;Lo;0;L;;;;;N;;;;;

1020A;NORTHERN PALAEOHISPANIC LETTER BA;Lo;0;L;;;;;N;;;;;

1020B;NORTHERN PALAEOHISPANIC LETTER BE;Lo;0;L;;;;;N;;;;;

1020C;NORTHERN PALAEOHISPANIC LETTER BI;Lo;0;L;;;;;N;;;;;

1020D;NORTHERN PALAEOHISPANIC LETTER BO;Lo;0;L;;;;;N;;;;;

1020E;NORTHERN PALAEOHISPANIC LETTER BU;Lo;0;L;;;;;N;;;;;

1020F;NORTHERN PALAEOHISPANIC LETTER KA WITH TWO ADDITIONAL STROKES;Lo;0;L;;;;;N;;;;;

10210;NORTHERN PALAEOHISPANIC LETTER KA WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10211;NORTHERN PALAEOHISPANIC LETTER KA;Lo;0;L;;;;;N;;;;;

10212;NORTHERN PALAEOHISPANIC LETTER KE WITH TWO ADDITIONAL STROKES;Lo;0;L;;;;;N;;;;;

10213;NORTHERN PALAEOHISPANIC LETTER KE WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10214;NORTHERN PALAEOHISPANIC LETTER KE;Lo;0;L;;;;;N;;;;;

10215;NORTHERN PALAEOHISPANIC LETTER KI WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10216;NORTHERN PALAEOHISPANIC LETTER KI;Lo;0;L;;;;;N;;;;;

10217;NORTHERN PALAEOHISPANIC LETTER KO WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10218;NORTHERN PALAEOHISPANIC LETTER KO;Lo;0;L;;;;;N;;;;;

10219;NORTHERN PALAEOHISPANIC LETTER KU WITH DOT;Lo;0;L;;;;;N;;;;;

1021A;NORTHERN PALAEOHISPANIC LETTER KU;Lo;0;L;;;;;N;;;;;

1021B;NORTHERN PALAEOHISPANIC LETTER TA WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

1021C;NORTHERN PALAEOHISPANIC LETTER TA;Lo;0;L;;;;;N;;;;;

1021D;NORTHERN PALAEOHISPANIC LETTER TE WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

1021E;NORTHERN PALAEOHISPANIC LETTER TE;Lo;0;L;;;;;N;;;;;

1021F;NORTHERN PALAEOHISPANIC LETTER TI WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10220;NORTHERN PALAEOHISPANIC LETTER TI;Lo;0;L;;;;;N;;;;;

10221;NORTHERN PALAEOHISPANIC LETTER TO WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10222;NORTHERN PALAEOHISPANIC LETTER TO;Lo;0;L;;;;;N;;;;;

10223;NORTHERN PALAEOHISPANIC LETTER TU WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10224;NORTHERN PALAEOHISPANIC LETTER TU;Lo;0;L;;;;;N;;;;;

10225;NORTHERN PALAEOHISPANIC LETTER L;Lo;0;L;;;;;N;;;;;

10226;NORTHERN PALAEOHISPANIC LETTER A2;Lo;0;L;;;;;N;;;;;

10227;NORTHERN PALAEOHISPANIC LETTER N2;Lo;0;L;;;;;N;;;;;

10228;NORTHERN PALAEOHISPANIC LETTER N1 WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10229;NORTHERN PALAEOHISPANIC LETTER N1;Lo;0;L;;;;;N;;;;;

1022A;NORTHERN PALAEOHISPANIC LETTER N3;Lo;0;L;;;;;N;;;;;

1022B;NORTHERN PALAEOHISPANIC LETTER N4;Lo;0;L;;;;;N;;;;;

1022C;NORTHERN PALAEOHISPANIC LETTER R1;Lo;0;L;;;;;N;;;;;

1022D;NORTHERN PALAEOHISPANIC LETTER R2 WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

1022E;NORTHERN PALAEOHISPANIC LETTER R2;Lo;0;L;;;;;N;;;;;

1022F;NORTHERN PALAEOHISPANIC LETTER S1 WITH ADDITIONAL STROKE;Lo;0;L;;;;;N;;;;;

10230;NORTHERN PALAEOHISPANIC LETTER S1;Lo;0;L;;;;;N;;;;;

10231;NORTHERN PALAEOHISPANIC LETTER S2;Lo;0;L;;;;;N;;;;;

10232;NORTHERN PALAEOHISPANIC LETTER S87;Lo;0;L;;;;;N;;;;;

All the NORTHERN PALAEOHISPANIC numerals and fractions from 10235 to 10239 are as below.

10233;NORTHERN PALAEOHISPANIC NUMERAL ONE;No;0;L;;;;1;N;;;;;

10234;NORTHERN PALAEOHISPANIC NUMERAL A;So;0;L;;;;;N;;;;;

10235;NORTHERN PALAEOHISPANIC NUMERAL B;So;0;L;;;;;N;;;;;

10236;NORTHERN PALAEOHISPANIC FRACTION ONE QUARTER;No;0;L;;;;1/4;N;;;;;

10237;NORTHERN PALAEOHISPANIC FRACTION ONE HALF;No;0;L;;;;1/2;N;;;;;

The word separator will be encoded as below.

10238;NORTHERN PALAEOHISPANIC SEPARATOR ;No;0;L;;;;1/2;N;;;;;

c	1020	1021	1022	1023	1024
0	A WITH ADDITIONAL STROKE	KA WITH ADDITIONAL STROKE	TI	S1	
1	A	KA	TO WITH ONE ADDITIONAL STROKE	S2	
2	E WITH ADDITIONAL STROKE	KE WITH TWO ADDITIONAL STROKES	TO	S87	
3	E	KE WITH ADDITIONAL STROKE	TU WITH ADDITIONAL STROKE	ONE	
4	I WITH ADDITIONAL STROKE	KE	TU	NUMERAL A	
5	I	KI WITH ADDITIONAL STROKE	L	NUMERAL B	
6	O WITH ADDITIONAL STROKE	KI	A2	FRACTION ONE QUARTER	
7	O	KO WITH ADDITIONAL STROKE	N2	FRACTION ONE HALF	
8	U WITH ADDITIONAL STROKE	KO	N1 WITH ADDITIONAL STROKE	WORD SEPARATOR	
9	U	KU WITH DOT	N1		
A	BA	KU	N3		
B	BE	TA WITH ADDITIONAL STROKE	N4		
C	BI	TA	R1		
D	BO	TE WITH ADDITIONAL STROKE	R2 WITH ADDITIONAL STROKE		
E	BU	TE	R2		
F	KA WITH TWO ADDITIONAL STROKES	TI WITH ONE ADDITIONAL STROKE	S1 WITH ADDITIONAL STROKE		

10. Proposed Characters

10200 P NORTHERN PALAEOHISPANIC LETTER A WITH ADDITIONAL STROKE

- Northeastern Iberian (Dual extended)

10201 D NORTHERN PALAEOHISPANIC LETTER A

10202 NORTHERN PALAEOHISPANIC LETTER E WITH ADDITIONAL STROKE

- Northeastern Iberian (Dual extended)

10203 NORTHERN PALAEOHISPANIC LETTER E

10204 NORTHERN PALAEOHISPANIC LETTER I WITH ADDITIONAL STROKE

- Northeastern Iberian (Dual extended)

10205 NORTHERN PALAEOHISPANIC LETTER I

10206 NORTHERN PALAEOHISPANIC LETTER O WITH ADDITIONAL STROKE

- Northeastern Iberian (Dual extended)

10207 NORTHERN PALAEOHISPANIC LETTER O

10208 NORTHERN PALAEOHISPANIC LETTER U WITH ADDITIONAL STROKE

- Northeastern Iberian (Dual extended)

10209 NORTHERN PALAEOHISPANIC LETTER U

1020A NORTHERN PALAEOHISPANIC LETTER BA

1020B NORTHERN PALAEOHISPANIC LETTER BE

1020C NORTHERN PALAEOHISPANIC LETTER BI

1020D NORTHERN PALAEOHISPANIC LETTER BO

1020E NORTHERN PALAEOHISPANIC LETTER BU

1020F NORTHERN PALAEOHISPANIC LETTER KA WITH TWO ADDITIONAL STROKES

10210 NORTHERN PALAEOHISPANIC LETTER KA WITH ADDITIONAL STROKE

10211 NORTHERN PALAEOHISPANIC LETTER KA

10212 NORTHERN PALAEOHISPANIC LETTER KE WITH TWO ADDITIONAL STROKES

10213 NORTHERN PALAEOHISPANIC LETTER KE WITH ADDITIONAL STROKE

10214 NORTHERN PALAEOHISPANIC LETTER KE

10215 NORTHERN PALAEOHISPANIC LETTER KI WITH ADDITIONAL STROKE

10216 NORTHERN PALAEOHISPANIC LETTER KI

10217 NORTHERN PALAEOHISPANIC LETTER KO WITH ADDITIONAL STROKE

10218 NORTHERN PALAEOHISPANIC LETTER KO

10219 NORTHERN PALAEOHISPANIC LETTER KU WITH DOT

1021A NORTHERN PALAEOHISPANIC LETTER KU

1021B 𐈛 NORTHERN PALAEOHISPANIC LETTER TA WITH ADDITIONAL STROKE

1021C 𐈜 NORTHERN PALAEOHISPANIC LETTER TA

1021D 𐈝 NORTHERN PALAEOHISPANIC LETTER TE WITH ADDITIONAL STROKE

1021E 𐈞 NORTHERN PALAEOHISPANIC LETTER TE

1021F 𐈟 NORTHERN PALAEOHISPANIC LETTER TI WITH ADDITIONAL STROKE

10220 𐈠 NORTHERN PALAEOHISPANIC LETTER TI

10221 𐈡 NORTHERN PALAEOHISPANIC LETTER TO WITH ADDITIONAL STROKE

10222 𐈢 NORTHERN PALAEOHISPANIC LETTER TO

10223 𐈣 NORTHERN PALAEOHISPANIC LETTER TU WITH ADDITIONAL STROKE

10224 𐈤 NORTHERN PALAEOHISPANIC LETTER TU

10225 𐈥 NORTHERN PALAEOHISPANIC LETTER L

10226 𐈦 NORTHERN PALAEOHISPANIC LETTER A2

10227 𐈧 NORTHERN PALAEOHISPANIC LETTER N2

10228 𐈨 NORTHERN PALAEOHISPANIC LETTER N1 WITH ADDITIONAL STROKE

10229 𐈩 NORTHERN PALAEOHISPANIC LETTER N1

1022A 𐈪 NORTHERN PALAEOHISPANIC LETTER N3

1022B 𐈫 NORTHERN PALAEOHISPANIC LETTER N4

1022C 𐈬 NORTHERN PALAEOHISPANIC LETTER R1

1022D 𐈭 NORTHERN PALAEOHISPANIC LETTER R2 WITH ADDITIONAL STROKE

1022E 𐈮 NORTHERN PALAEOHISPANIC LETTER R2

1022F 𐈯 NORTHERN PALAEOHISPANIC LETTER S1 WITH ADDITIONAL STROKE

10230 𐈰 NORTHERN PALAEOHISPANIC LETTER S1

10231 𐈱 NORTHERN PALAEOHISPANIC LETTER S2

10232 𐈲 NORTHERN PALAEOHISPANIC LETTER S87

10233 𐈳 NORTHERN PALAEOHISPANIC NUMERAL ONE

10234 𐈴 NORTHERN PALAEOHISPANIC NUMERAL A

10235 𐈵 NORTHERN PALAEOHISPANIC NUMERAL B

10236 𐈶 NORTHERN PALAEOHISPANIC FRACTION ONE QUARTER

10237 𐈷 NORTHERN PALAEOHISPANIC FRACTION ONE HALF

10238 𐈸 NORTHERN PALAEOHISPANIC SEPARATOR

11. Glyph Variation Chart

Encoded glyph		Transcription (see pp. 4-5)	Northern						
			Iberian			Celtiberian			
			NI+	NI++	NI-	WC+	WC-	EC+	EC-
	10200	á							
	10201	a							
	10202	é							
	10203	e							
	10204	í							
	10205	i							
	10206	ó							
	10207	o							
	10208	ú							
	10209	u							
	1020A	ba							
	1020B	be							
	1020C	bi							
	1020D	bo							

Code	Reading							
1020E	**bu**							
1020F	**ḱa**							
10210	**ka**							
10211	**ga**		**ka/ga**		**ka/ga**	**ka/ga**			**ka/ga**
10212	**ḱe**							
10213	**ke**							**ke/ge**
10214	**ge**		**ke/ge**		**ke/ge**			
10215	**ki**							
10216	**gi**		**ki/gi**		**ki/gi**			**ki/gi**
10217	**ko**							
10218	**go**		**ko/go**		**ko/go**			**ko/go**
10219	**ku**		**ku/gu**		**ku/gu**			**ku/gu**
1021A	**gu**							
1021B	**ta**							
1021C	**da**		**ta/da**		**ta/da**			**ta/da**
1021D	**te**							

	Code	Value							
⊖	1021E	**de**							
	1021F	**ti**							
	10220	**di**			**ti/di**	**ti**	**ti/di**		**ti/di**
	10221	**to**							
	10222	**do**			**to/do**	**to**	**to/do**		**to/do**
	10223	**tu**					**tu/du**		
	10224	**du**			**tu/du**				**tu/du**
	10225	**l**							
	10226	**â**							
	10227	**m**							
	10228	**ń**							
	10229	**n**				**m**	**m**		
	1022A	**ḿ**				**n**	**n**		
	1022B	**m̌**				**?**	**?**	**?**	**?**
	1022C	**r**							
	1022D	**ŕ**				**r**		**r**	

15

	Code	Name							
	1022E	ř			ŕ		r		r
	1022F	ŝ	s			z		z	
	10230	s				z	z		z
	10231	ś	s		s	s	s	s	s
	10232	S87							
	10233	numeral 1							
	10234	numeralA							
	10235	numeralB							
	10236	Fraction 1/4							
	10237	Fraction 1/2							
	10238	Separator							

NI+ = Northeastern Iberian dual standard

NI++ = Northeastern Iberian dual extended

NI- = Northeastern Iberian non dual

WC+ = Western Celtiberian dual

WC- = Western Celtiberian non dual

EC+ = Eastern Celtiberian dual

EC- = Eastern Celtiberian non dual

12.- Images.

Phoenician	'	b	g	d	h	w	z	ḥ	ṭ	y	k	l	m	n	s	'	p	ṣ	q	r	š	t						
Southern PH																												
Northern PH																												

Fig. 1.- Comparison between Phoenician, Southern Palaeohispanic and Northern Palaeohispanic.
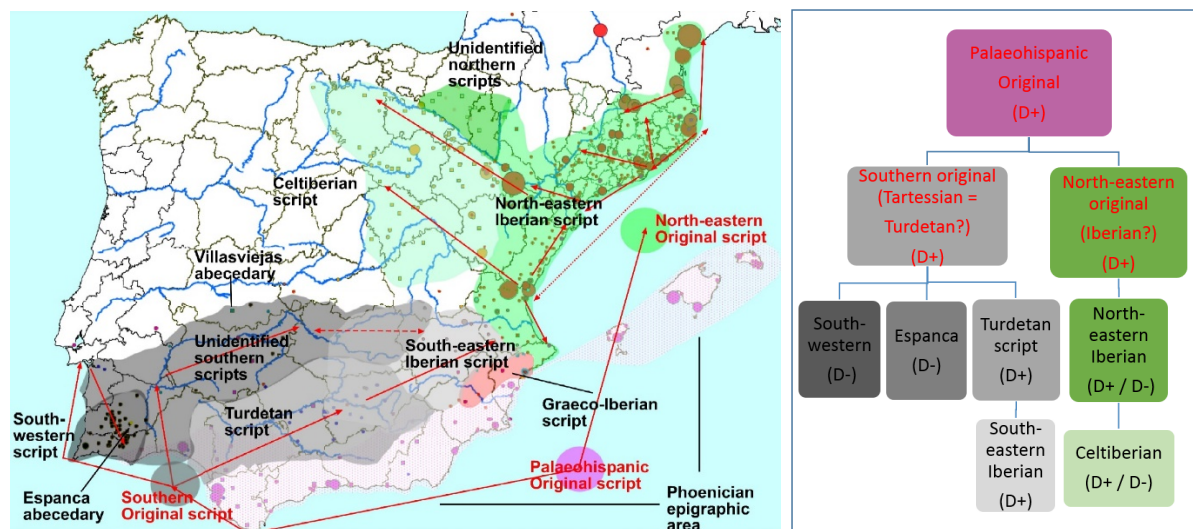
Fig. 2.- Palaeohispanic scripts. Most recent proposal of diffusion and Genealogic model (Ferrer i Jané 2018).

Northeastern Iberian dual extended

10206
10225
1022F
10200
10221
10207
10230
10201
10226
10222
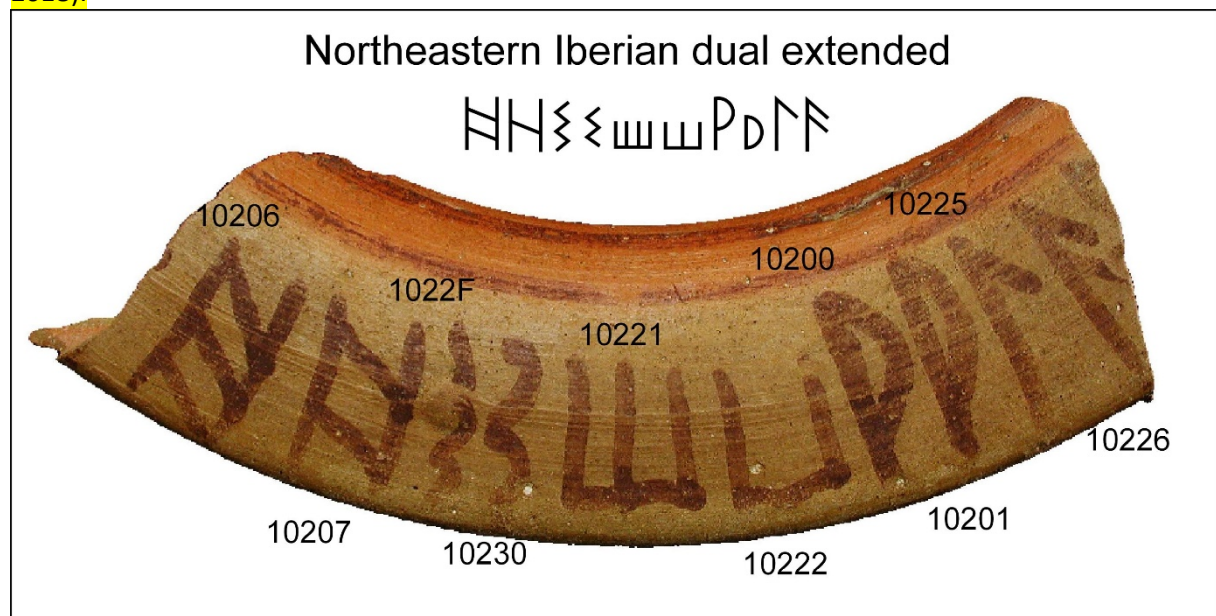
Fig. 3.- Example showing an inscription with code points and expected letter shapes in online/print publications (Castellet de Bernabé: fragment of a dual extended abecedary)[3].

---

[3] The two signs of the duality of the vowel **a** are apparently equal, but, it does not make any sense, as it is a dual abecedary. We suppose that the second a is drew in a clumsy way but is in fact the sign D.
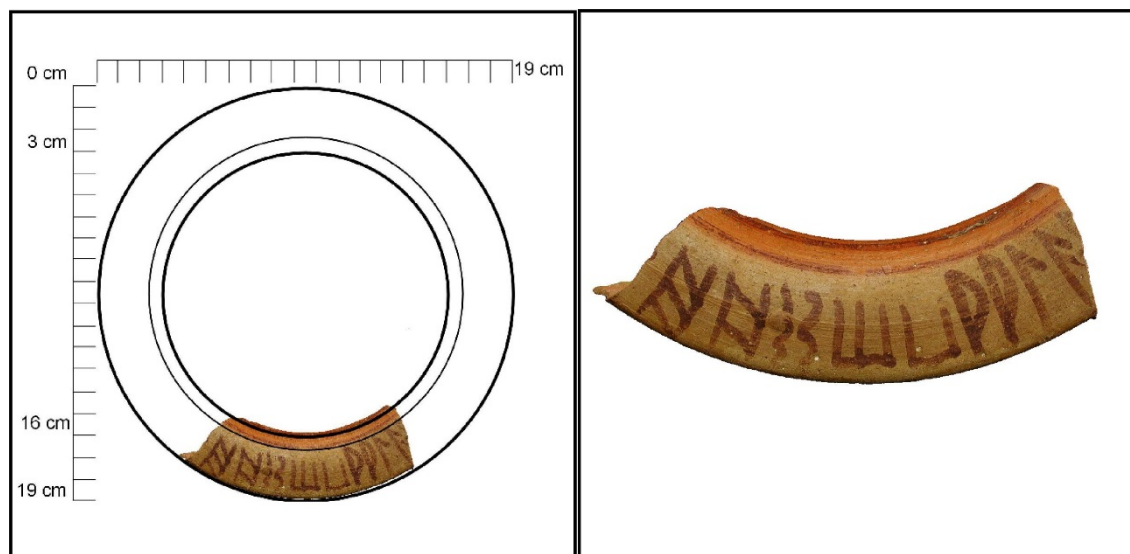
Fig 4.-Castellet de Bernabé's abecedary (Extended dual script). <mark>See fig. 3 for its codification.</mark>
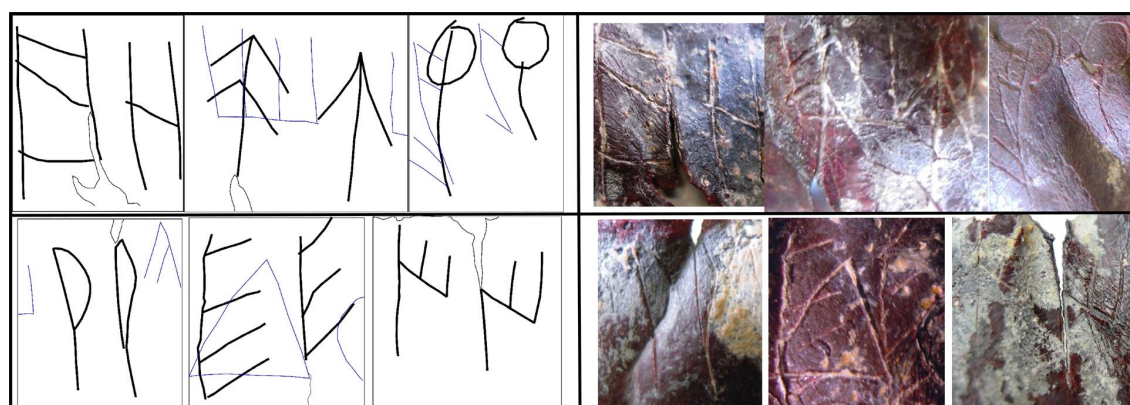
Fig 5.- Tos Pelat's abecedary (F.13.77*) (Extended dual script). Above, detail of dualities for vowels[4] and trill. Below: General view.
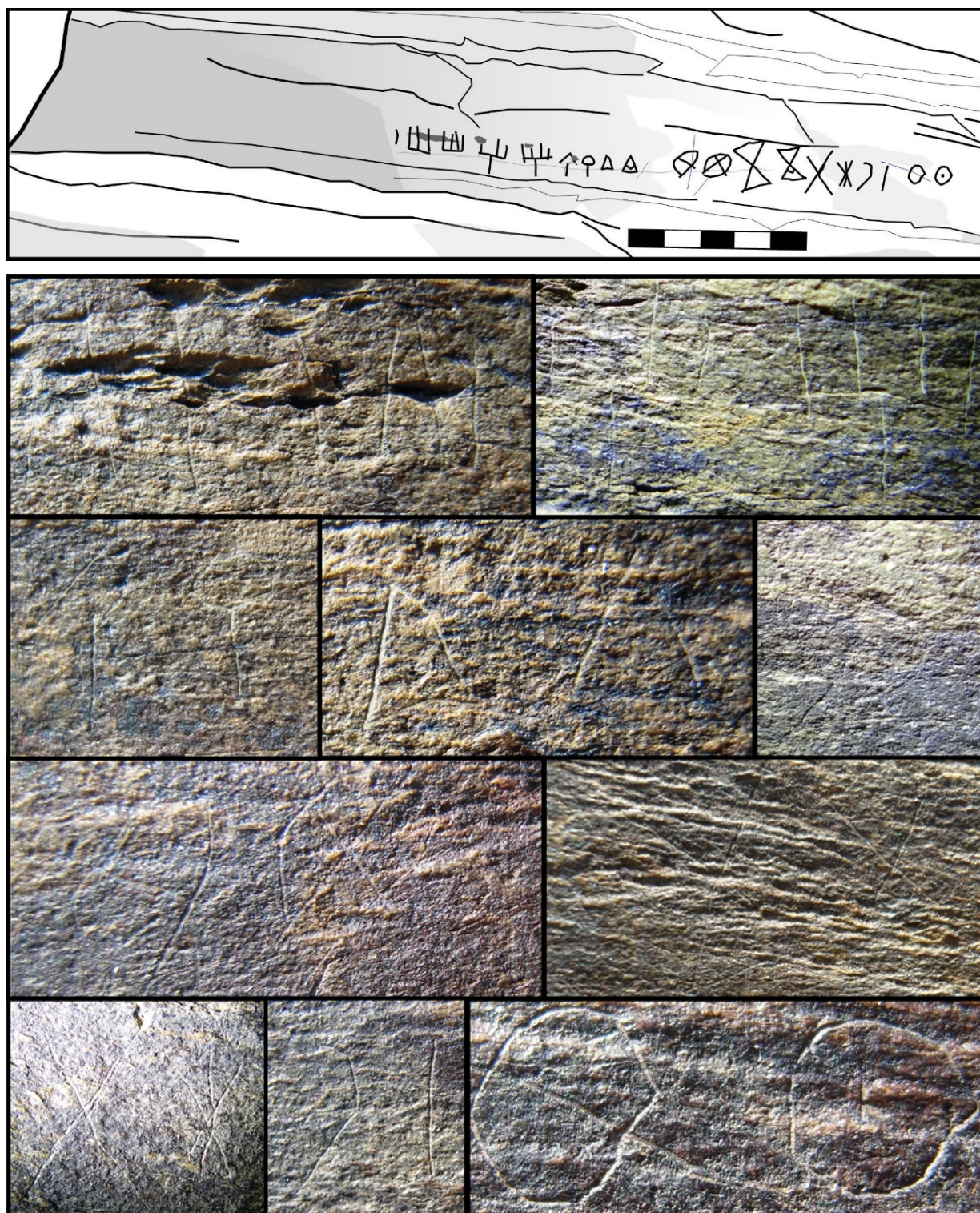
Fig 6.-Bolvir's abecedary (Standard dual script). Above: general view. Below: Detailed pictures.
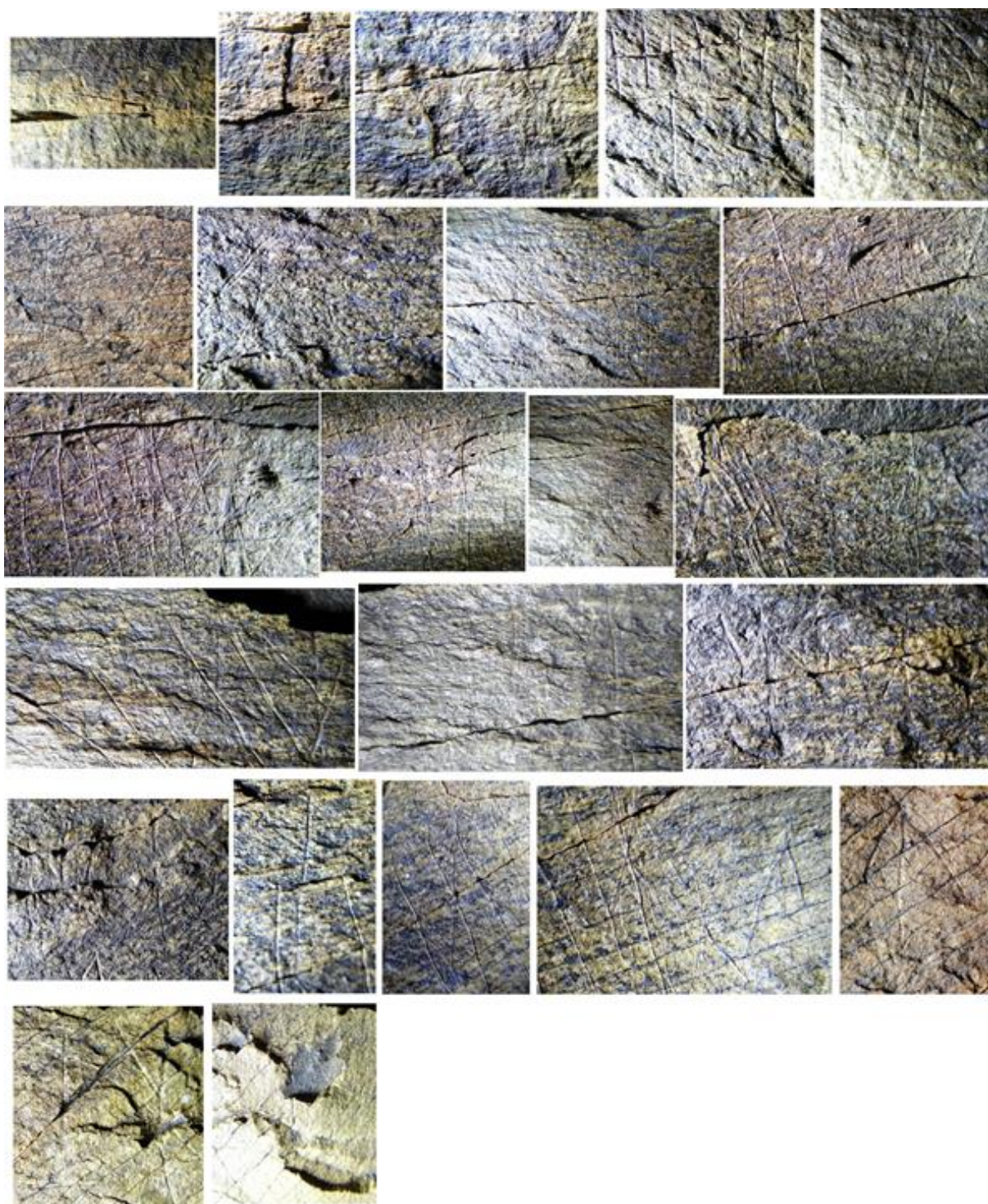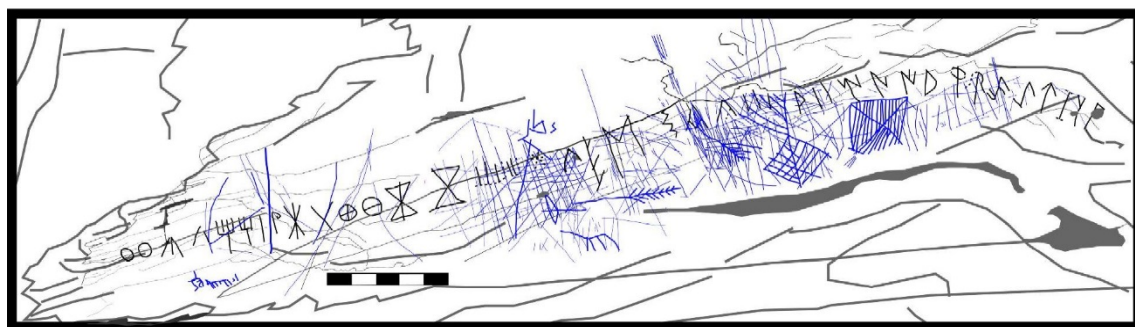
Fig 7.-Ger's abecedary (Standard dual script).
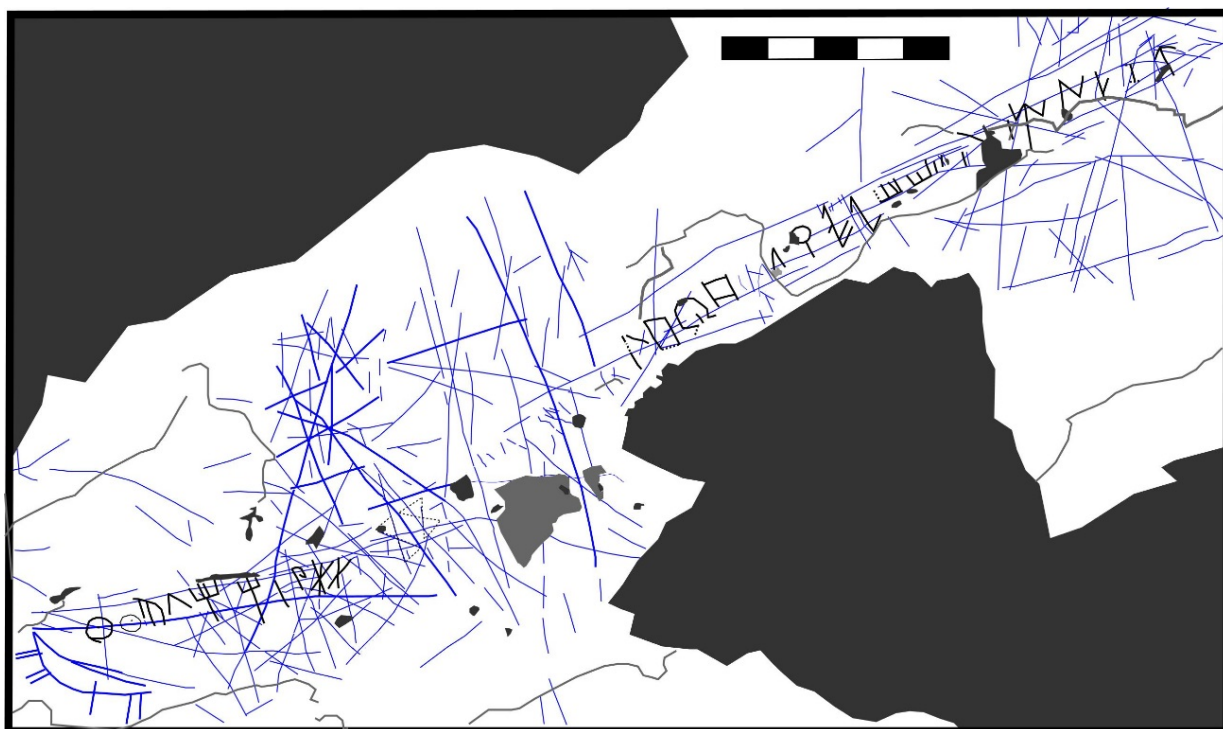
Fig 8.-La Tor de Querol's abecedary (Standard dual script).

Fig. 9.-Simplified abecedaries from Can Rodon (Non-dual script).
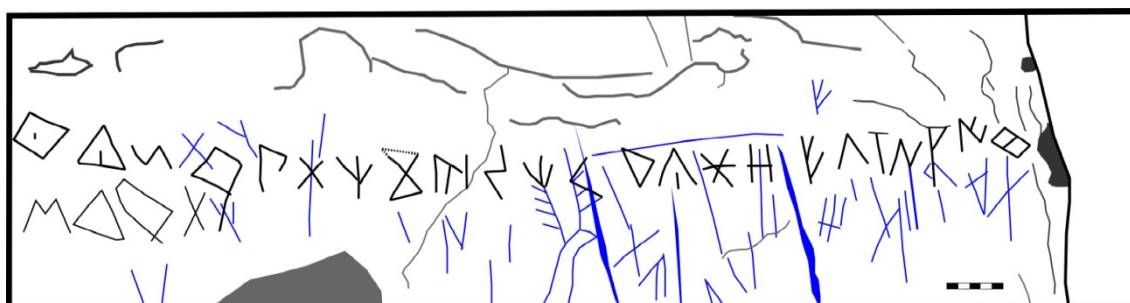
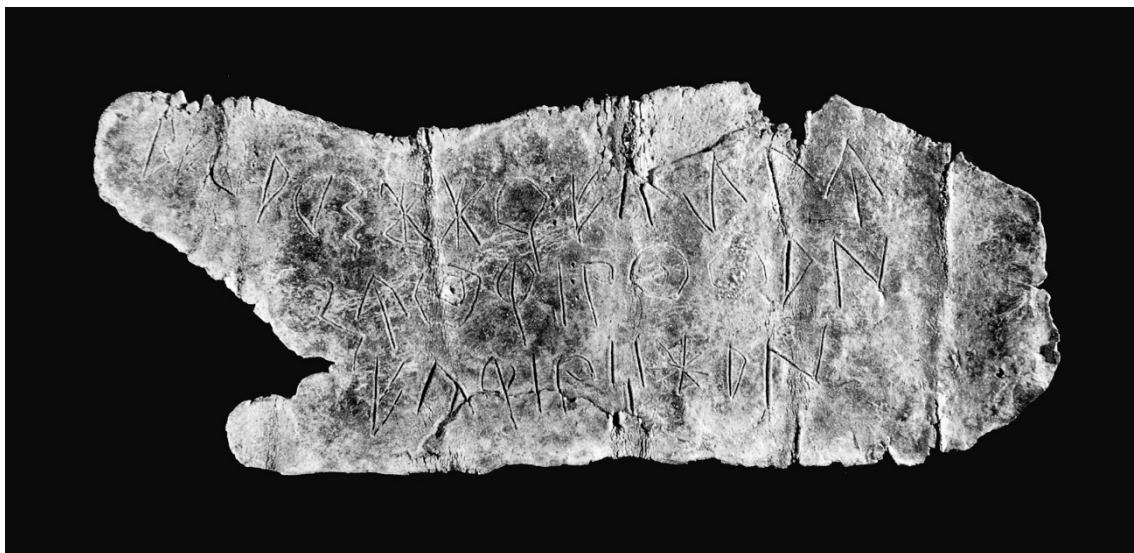**Fig. 10.-**Esquirol's abecedary (Non-dual script).

Fig 11.- Lead sheet from Ullastret (C.2.4) (dual script).

Fig 12.- Lead sheet from La Balaguera (F.17.*) (dual script).

Fig 13.- Lead sheet from Castellet de Bernabé (F.13.75*). Extended dual abecedary with explicit duality for **ŕ**.

Fig 14.-Stone plaque from Empúries (non-dual script).

Fig 15.- Stele from Vispesa bearing a non-dual northeastern Iberian inscription.

Fig 16.- Bronze coin from **undikesken** (non-dual script), with mark of value: e= (1/2)

Fig 17.-Bronze coin from **undikesken** (non-dual script), with mark of value: e- (1/4)

Fig 18.- Ceramic vase from Terrassa bearing the name **talskubilos** (non-dual script).

Fig 19.- Spindle-whorl from Gebut (non-dual script, right to left).

Fig 20.-Painted inscription in a ceramic vase from Llíria (F.13.5) (Extended dual script).

Fig 21.- Stamps on dolium from Pech Maho (B.7.32) (Standard dual script).

Fig 22.- Vase from Joncosa (D.18.1*) (non-dual script).

Fig 23.- Lead plaque from Yátova (F.20.1) (non-dual script) with numerals (red circle).

Fig 24.- Northeastern Iberian script (de Hoz 2011).

13. Bibliography

BURRIEL, J. Mª, MATA, C., FERRER I JANÉ, J, RUIZ, A.L. VELAZA, J., PEIRÓ, Mª A., ROLDÁN, C., MURCIA, S., DOMÉNECH, A. (2011): "El plomo escrito del Tos Pelat (Moncada, Valencia)", *Palaeohispanica* 11, 191-224.

CAMPMAJO, P. UNTERMANN, J. (1991): "Corpus des gravures ibériques de Cerdagne", *Ceretania* 1, 39-59.

CORREA, J.A. (1992): "Representación gráfica de la oposición de sonoridad en las oclusivas ibéricas (semisilabario levantino)", *AION* 14, 253-292.

FERRER I JANÉ, J. (2005): "Novetats sobre el sistema dual de diferenciació gràfica de les oclusives", *Palaeohispanica 5*, 957-982.

FERRER I JANÉ, J. (2007): "Sistemes de marques de valor lèxiques sobre monedes ibèriques", *Acta Numismàtica* 37, 53-73.

FERRER I JANÉ, J. (2011b): "Sistemas metrológicos en textos ibéricos (1): del cuenco de La Granjuela al plomo de La Bastida". *E.L.E.A.* 11, 99-130.

FERRER I JANÉ, J. (2013a): "Deux alphabets duals rupestres de Cerdagne", *Sources – Les cahiers de l'Âne Rouge* 1, 9-18.

FERRER I JANÉ, J. (2013b): "Els sistemes duals de les escriptures ibèriques", *Palaeohispanica* 13, 445-459.

FERRER I JANÉ, J. (2014): "A propòsit d'un sisè de bronze de baitolo amb la llegenda be II", *Acta Numismàtica* 44, 57-69.

FERRER I JANÉ, J. (2014b): "Ibèric kutu i els abecedaris ibèrics", *Veleia* 31, 227-259.

FERRER I JANÉ, J. (2014c): "Deux noveaux alphabets ibères rupestres de Cerdagne", *Sources – Les cahiers de l'Âne Rouge* 2, 11-20.

FERRER I JANÉ, J. (2015): "Las dualidades secundarias de la escritura ibérica nororiental", *ELEA* 14, 309-364.

FERRER I JANÉ, J. (2017a): "El origen dual de las escrituras paleohispánicas: un nuevo modelo genealógico", *Palaeohispanica* 17, 55-94.

FERRER I JANÉ, J. (2018a): *L'origen i el desenvolupament de les escriptures paleohispàniques*. U.B. Tesi Doctoral.

FERRER I JANÉ, J. (2019): "A la recerca del trial: les variants supercomplexes de les escriptures paleohispàniques", *Actas del XV Congreso de Estudios Clásicos de la SEEC (Valladolid, July 2019) = Palaeohispanica* 19, 27-53.

FERRER I JANÉ, J. Y MONCUNILL, N. (2019): "Chapter 4. Palaeohispanic writing systems: classification, origin and development", *Palaeohispanic Languages and Epigraphies*, Oxford University Press, 78-108.

FERRER I JANÉ, J., SÁNCHEZ, M. (2017): "L'enigma b'oide al descobert: **kaśtaum** i **baikar** en sengles inscripcions ibèriques sobre una tortera i un vaset de Vilademuls", *RAP* 27, 221-236.

GÓMEZ-MORENO, M. (1922): "De epigrafía ibérica: El plomo de Alcoy", Revista de Filología Española, Tomo IX, 342-366.

GÓMEZ-MORENO, M. (1949): *Misceláneas, Historia, Arte, Arqueología*, Madrid 1949.

HOZ J. de (1985): "El nuevo plomo inscrito de Castell y el problema de las oposiciones de sonoridad en ibérico", *Symbolae Ludouico Mitxelena septuagenario oblatae*, 443-453.

HOZ, J. de, (1981): "Algunas precisiones sobre textos metrológicos ibéricos", *APL* 16, 475-486.

HOZ, J. DE (2010): *Historia lingüística de la Península Ibérica en la antigüedad. I. Preliminares y mundo meridional prerromano*. Madrid: Consejo Superior de Investigaciones Científicas.

HOZ, J. DE (2011): *Historia lingüística de la Península Ibérica en la antigüedad. II. El mundo ibérico preromano y la indoeuropeización*. Madrid: Consejo Superior de Investigaciones Científicas.

JORDÁN, C. (2005). ¿Sistema dual de escritura en celtibérico?, *Actas del IX Coloquio Internacional de Lenguas y Culturas Paleohispánicas (Barcelona 2004), Acta Palaeohispanica IX, Palaeohispanica*, 5: 1013-1030.

JORDÁN, C. (2007): Estudios sobre el sistema dual de escritura en epigrafía no monetal celtibérica, *Palaeohispanica*, 7: 101-142.

LEJEUNE, M. (1983): "Vieille-Toulouse et la métrologie ibèriques", *Revue Archeologique de Narbonnaise*, 16, 29-37.

MALUQUER, J. (1968): *Epigrafía prelatina de la península ibérica*, Barcelona.

MONCUNILL, N. (2007): *Lèxic d'inscripcions ibèriques (1991-2006),* Tesi Doctoral, Barcelona, Universitat de Barcelona.

ORDUÑA, E. (2005): "Sobre algunos posibles numerales en textos ibéricos", *Palaeohispanica* 5, 491-506.

ORDUÑA, E., (2006): *Segmentación de textos ibéricos y distribución de los segmentos*, Madrid, Departamento de Filología Clásica. Facultad de Filología de la UNED.

ORDUÑA, E. (2013): "Los numerales ibéricos y el vascoiberismo", *Palaeohispanica* 13, 517-529.

PANOSA, M.I. (1993): "Nuevas inscripciones ibéricas de Cataluña", *Complutum* 4, 175-222.

QUINTANILLA, A. (1992): "Sobre la notación en la escritura ibérica del modo de articulación de las consonantes oclusivas", *Studia Palaeohispanica et Indogermánica J. Untermann ab Amicis Hispanicis Oblata*, *Aura Saecula* 10, Universitat de Barcelona, 239-250.

RODRÍGUEZ RAMOS, J. (2001): "Signos de lectura problemática en la escritura ibérica", *Archivo Español de Arqueología* 74, Madrid, 281-290.

RODRÍGUEZ RAMOS, J. (2004): *Análisis de epigrafía íbera,* Anejos de Veleia, Series Minor 12, Vitoria-Gasteiz, Universidad del País Vasco.

RODRÍGUEZ RAMOS, J. (2005): "Introducció a l'estudi de les inscripcions ibèriques", Revista de la Fundació Privada Catalana per l'Arqueologia ibèrica 1, 13-144.

SIMON, I. (2011): "Interpunctiones palaeohispanicae", Epigraphica LXXIII, 1-2, 87-108.

UNTERMANN, J. (1975): *Monumenta Linguarum Hispanicarum. I Die Münzlegenden*, Wiesbaden.

UNTERMANN, J. (1980): *Monumenta Linguarum Hispanicarum. II Die iberischen Inschriften aus Sudfrankreicht*c

UNTERMANN, J. (1990): *Monumenta Linguarum Hispanicarum Band III. Die iberischen Inschriften aus Spanien*, Wiesbaden.

VELAZA, J. (1991): *Léxico de las inscripciones ibéricas: 1976-1989,* Universitat de Barcelona, Barcelona.

VELAZA, J. (2006): "Chronica epigraphica iberica VII (2004-2005)", *Palaeohispanica* 6, 2006, 303-327.

VELAZA, J. (2009): "Epigrafía y literacy paleohispánicas en territorio vascón: Notas para un balance provisional", *Palaeohispanica* 9, 61-622.

14. Acknowledgement

A. Administrative

1. **Title:**	***Proposal to encode the NORTHERN PALAEOHISPANIC script***
2. Requester's name:	*Joan Ferrer, Noemí Moncunill, Javier Velaza, and Deborah Anderson*
3. Requester type (Member body/Liaison/Individual contribution):	*liaison*
4. Submission date:	*FILL IN*
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	*x*
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	*x*
Proposed name of script:	*NORTHERN PALAEOHISPANIC*
b. The proposal is for addition of character(s) to an existing block:	
Name of the existing block:	
2. Number of characters in proposal:	*58*
3. Proposed category (select one from below - see section 2.2 of P&P document):	

A-Contemporary		B.1-Specialized (small collection)		B.2-Specialized (large collection)	
C-Major extinct		D-Attested extinct	*x*	E-Minor extinct	
F-Archaic Hieroglyphic or Ideographic		G-Obscure or questionable usage symbols			

4. Is a repertoire including character names provided?	*yes*
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	*yes*
b. Are the character shapes attached in a legible form suitable for review?	*yes*

---

[5] Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

The Atelier National de Recherche Typographique (Thomas Huot-Marchand and Arthur Francietta)

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

Fill in

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*

b. Are published examples of use (such as samples from newspapers, magazines, or other sources)

of proposed characters attached? *yes*

7. Special encoding issues:

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. Additional Information:

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/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *yes*

   If YES explain *L2/15-119, L2/15-120, and L2/15-012*

2. Has contact been made to members of the user community (for example: National Body,

   user groups of the script or characters, other experts, etc.)? *yes*

   If YES, with whom? *Scholars in Spain, Portugal, France, Germany, and UK*

   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? *yes*

   Reference: *See proposal*

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

   Reference: *See proposal*

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

   If YES, where? Reference: *Books, articles, database, etc.*

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

   in the BMP? *no*

   If YES, is a rationale provided?

   If YES, reference:

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

8. Can any of the proposed characters be considered a presentation form of an existing

   character or character sequence?

   If YES, is a rationale for its inclusion provided? *yes*

   If YES, reference: *See proposal*

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

   existing characters or other proposed characters? *no*

   If YES, is a rationale for its inclusion provided?

   If YES, reference:

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

    to, or could be confused with, an existing character? *yes*

    If YES, is a rationale for its inclusion provided? *yes*

    If YES, reference: *See proposal*

11. Does the proposal include use of combining characters and/or use of composite sequences? *no*

If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	*no*
If YES, reference:	

12. Does the proposal contain characters with any special properties such as

control function or similar semantics?　　　　　　　　　　　　　　　　*no*

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

13. Does the proposal contain any Ideographic compatibility characters?　　*no*

　　If YES, are the equivalent corresponding unified ideographic characters identified?

　　　　If YES, reference: