

To: Unicode Technical Committee
From: Debbie Anderson, SEI, UC Berkeley
Date: October 8, 2018
Subject: Comments and initial review of L2/17-337

Background

Duncan Poupard, Assistant Professor in the Dept. of Translation, Chinese University of Hong Kong, inquired about the status of Naxi Dongba. He is a researcher of the Naxi and is working on digitizing the 1932 Naxi Bible translation. Ideally, he would like to have the Gospel of Mark in the modern Dongba script.

He reports:

The Naxi script (more prominently dongba script, geba is very rare) is still being used in some limited contexts in the Lijiang area, and thanks to government initiatives there are a number of students learning it, not to mention a relatively recent law stating that all shop fronts in the Old Town area of Lijiang must have a dongba translation...

There has been a bit of confusion over what exactly the Naxi dongba script is in the west, but I believe publications such as Popular Dongba Script (通俗东巴文) by He Limin 和力民 (2007) and research by Yu Suisheng and his team at Southwest University in Chongqing, show that dongba can be used not as a mnemonic script but with a near one-to-one representation of graph to phoneme. My current research (<http://journal.oraltradition.org/issues/32i/poupard>) is also an attempt at re-evaluating the script....

I was slightly disappointed to see the China side shut down the inclusion of the geba syllabary: in vernacular dongba writing there are plenty of examples where a geba character is used to represent a grammatical particle to avoid writing out a complicated Naxi dongba rebus.... the scripts are not as separate as all that - and in modern usage I believe they could take on role similar to Japanese katakana in transcribing foreign names etc. That said, I don't want the process to get slowed down any further, and the geba script can always be included at a later date.

Prof. Poupard offered to review [L2/17-337](#) (Results of the ad-hoc meeting on Naxi Dongba in Hohhot [WG2 N4895]). Below are some preliminary comments.

[Comments from Debbie Anderson]

The repertoire in [L2/17-337](#) needs to be checked; with a repertoire of 1188 characters, it is possible there may be errors.

[Duncan Poupard]

The repertoire is smaller than those in earlier proposals, and indeed represents fewer graphs than appear in Fang Guoyu's dictionary (and certainly Joseph Rock's).

I believe the reduction can largely be explained by the omission of proper names of people and places, and wonder if this was purposeful¹ – there are plenty of distinct compound graphs for names of places in the Naxi areas and of course for all kinds of heroes, gods and famous dongba practitioners. This does mean that very commonly seen compound graphs



such as the one for Lijiang itself, *Yi gv ddiuq* , are not present in the current proposal. To write these names with the currently proposed character set is possible, but one must use the separate graphs for *yi*, *gv* and *ddiuq* in a syllabic fashion, for example.



The graph for *ngv lv* (Fang No. 109), snow mountain, used almost exclusively to refer to Lijiang's Jade Dragon snow mountain, is not present in the current proposal, suggesting that the word would have to be written with the two phonetic elements, "silver"



and "stone" : (ngvq, 1AAFA; lv, 1AA90).

This is the drawback to a cut-down character set: you may be able to recreate words phonetically via combining syllables, but the original *logographs* are in danger of being lost. It would be a shame to move from a logography (although not all Naxi graphs are logographic, many are) to a syllabary because of Unicode encryption. I would recommend a thorough read through of the list alongside a base set (I think Fang Guoyu's dictionary would be fine for this) and note any important omissions. I'd be happy to do this.

[Debbie Anderson]

The names of the characters should also be double-checked.

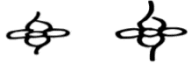
[Duncan Poupard]


The Naxi pinyin used looks on the whole very well done, but does need to be proofed.

Some graphs do have multiple readings. *Oq* (jade/turquoise, 1AB25) and *herq* (green, 1A9C2) are the same graph, and appear as two distinct entries (with slightly different calligraphy) in the proposal.²


¹ D. Poupard writes in a later email, "I believe the reason for the slightly reduced graph inventory may be a drive at simplification for modern usage, although as I say I'm wary of the script being turned into a syllabary."

² From a later email exchange: [Debbie] Do you think the distinction above is artificial, and these are actually the same glyph (and character)? [D. Poupard] For the jade/green graphs, they should be the same (that is, one graph, two separate readings, not two separate graphs, the distinction at present seems artificial to me).






The graph  (1ABFC), *Tei ee bbvq ddee*, conflates two variant pronunciations as one (it's either *tei'ee* or *bbvq ddee*).

I can also see that there are graphs with multiple readings where one reading is missing from the proposal.


Mi ( fire, 1AACD) and *hiuq* (red, Fang No. 1183) are represented by the same graph (also operating metaphorically like the above), but the proposal only contains *mi* (fire), not *hiuq*.



Importantly, the graph for *do bbaq* , (dongba) is not present, meaning that the

word would have to be encoded with two phonetic loans:   a wooden board (*do*, 1A907), and a flower (*bbaq*, 1A823). However, *biuq*, a variant reading of *do bbaq*, is

present -  (1A880), meaning that if you wanted to write '*do bbaq*' you could, but using the graph listed as '*biuq*'. Fang lists both readings in his dictionary.³

Rock lists two ways of writing "dongba". The first is the more

pictographic graph , read *do bbaq*, representing a dongba ritualist wearing the five-lobed crown (only three are visible, as it is a front-on view). Rock does also give a second method, using two rebuses, the first being the wooden board, the second being a

man suffering from a goitre, also pronounced *bba*   (this second graph is not present in the proposal).

Nevertheless, the number of polyphonic graphs is limited and as such I don't see using the names based on pronunciation as a problem.



³ From a later email exchange [D. Poupard]: As a lot of people are going to want to write 'dongba' (do bbaq) as it's the name of the script after all, this word should definitely have its own graph in my opinion (rather than two syllabic graphs).

[Debbie Anderson]


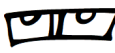
As noted in [L2/17-337](#), Naxi Dongba experts need to describe how the characters combine (i.e., which ones may appear on top of another, etc.).

[Duncan Poupard]

A full description of the composition of the Naxi graphs would be a time-consuming undertaking. There are few “rules” as such, but the graphs can be split into various categories.

Mai , an animal’s tail, can be combined with the graph for water, *jjj*, to make *jjj mai* , downstream, the “tail-end” of a river. This can be classified as a compound ideogram.

Plenty of graphs are combined with others and used for their phonetic value, the modern-

day Yi ethnic group are known as *lo lo*  (1AA86), and written with a plowtail ‘*loq*’  as a phonetic component above the head of a person. This would be a phono-semantic graph. The plowtail can, of course, be used independently (see Fang No. 849, not present in the proposal).

Certain graphic elements can also be used with existing graphs to change their meaning in predictable ways, such as the black dot to indicate something is black in colour⁴:



(black mountain, black cliff, black stone, black cloth – all present in Fang but not in the proposal).

Similarly, many dots can indicate the idea of plurality:

⁴D. Poupard writes in a later email: “I have found in an old unpublished dictionary from the early twentieth century an entry for the black dot used individually and given the explanation of “black, when this round dot is used above a graph it indicates that object is black”. So yes I agree it could indeed be included individually in the repertoire. It should have the reading ‘*naq*’.” See image in Appendix.



(a heaven full of stars, a land full of grass, flour – all present in the proposal).

While including every combinatory possibility would be difficult, especially in a limited character set, it is important to make sure that the basic component parts of the compound graphs that are present in the proposal are *also* included individually – such as the plowtail example above. This would require another close reading of the repertoire, but most of the more glaring omissions could be spotted by carrying out the comparison with Fang.

Appendix: Image showing black dot.

