

## Article

### *Assessing Laurent Sagart's Approach of Sino-Austronesian*

Arnaud Fournet

*Abstract:* The paper assesses the Sino-Austronesian family proposed by Laurent Sagart. This idea raises several issues. The first point is that there exists no reconstruction of Old Chinese, that would account for all Chinese dialects. What exists are reconstructions of Pre-Proto-Mandarin, but they cannot account for a number of divergent dialects, especially Min. What is more, it is obvious that the “reconstruction” of “Old Chinese”, used by Sagart in his comparison with Austronesian, seriously fails to account for Chinese dialects in the first place. Several words which Sagart reconstructs with the same word-final *\*-a(?)* have neither the same tone nor the same vowel nucleus in present-day dialects, and quite certainly none of them had *\*-a(?)* as main vowel in Old Chinese.

*Keywords:* Chinese, Austronesian, Sino-Tibetan, Sound correspondence.

#### 1. *Old Chinese and Proto-Mandarin*

As defined in Baxter's *Handbook of Old Chinese Phonology* (1992:1), *Old Chinese* is to be taken as “the ancestor of all attested varieties of Chinese, and the earliest stage of the Chinese language that can be reconstructed from Chinese alone”. An obvious and implicit premise underlying that definition is that *all the Chinese dialects* indeed originate in *one* proto-language and that they make up a valid *monophylogenetic* node. This premise has never been questioned so far. From a sheer technical and descriptive point of view, the so-called Chinese “dialects” are separate idioms, often very divergent from one another and most often mutually understandable. But the cultural history of China and the ethnolinguistic consequences of the political unification two thousand years ago have created the general feeling among Chinese people that these “dialects”, whatever their differences may be, are “constituting a single language”, as Baxter reminds us (1992:7).

In his book, Baxter claims to have reconstructed *Old Chinese*. It is more accurate to state that he has reconstructed *Proto-Mandarin* (PM) and *Pre-Proto-Mandarin* (PPM). The reasons for this unusual terminology will gradually become clear to the reader. Proto-Mandarin (Baxter's *Middle Chinese*) is the literary and politically dominant language spoken at the time of the Tang Dynasty. This language is the source of the numerous loanwords existing in Korean, Japanese and Vietnamese. PM has been phonetically described in the QieYun (601 AD), a work which analyses each character of the Chinese script in terms of Initials and Rhymes. This method is adapted from the alphabetic tradition with which Chinese people came in contact when Buddhism entered China. Although it remains indirect, a complete reconstruction of PM is possible on that basis and Baxter's *Handbook* is in our opinion one of the most consistent and complete reconstructions of PM available, if not the best one. As the QieYun takes into account some dialectal variations of the time it was written, PM as reconstructed with the QieYun accounts for a stage of Mandarin that is in fact slightly older than 601 AD.

The next thing that Baxter's handbook provides is a tentative reconstruction of PPM (Baxter's *Old Chinese*), that is to say an extrapolation of PM that accounts for the rhyming patterns existing in the ShuJing and the phonetic components embedded in the graphic system. PPM can be approximately dated back to -1000 BC.

## 2. *Reasons why Pre-Proto-Mandarin is not Old Chinese*

Baxter claims that his *Pre-Proto-Mandarin* reconstruction can be equated with *Old Chinese*. But it is quite clear that his PPM does not account for some of the Chinese dialects, and most notably, it does not account for Min Dialects, a set of extremely and notoriously divergent dialects spoken in the south-eastern province of FuJian. As a matter of fact, he never mentions Min data, which would indeed instantly expose the inadequacy of the equation  $PPM = OC$ . The names *Old Chinese* and *Middle Chinese* as used by Baxter are a rewording of Karlgren's *Archaic Chinese* and *Old Chinese*. This rewording represents the advances made since the time of Karlgren. In our opinion, Baxter's goals at the time when he wrote his Handbook were to build two consistent systems accounting for the QieYun and the ShuJing in the first place. Unfortunately, the names given to the reconstructions are highly misleading for the unexperienced readers. The same unfortunate situation exists with other works, such as Sagart's *The Roots of Old Chinese*. In fact, these (inadequate) names and rewordings are misleading for the authors themselves. What current and past sinologists call “Old Chinese” (our PPM) is related to the real Old Chinese proto-language in the same way as Mycenaean Greek is related to Proto-Indo-European. Nobody involved in Indo-European Studies would confuse Mycenaean Greek with PIE, but the fact is most sinologists confuse *Pre-Proto-Mandarin* with *Old Chinese*.

Because of this entrenched terminological confusion, no critical assessment of the ability of PPM to be equated with Old Chinese has ever been made. It is obvious that PPM does not account for Min dialects. And it must be underlined that this failure is not a minor issue which could be contained by creating a cordoning-off around a limited set of lexical items. The failure is extensive, massive, structural. It is about impossible to deal with any single word in any single Min dialect without bumping into considerable problems, as will be shown below. This unfortunate situation is hardly surprising as Min dialects have never been dealt with in another way than resorting to the always available remedial patch of so-called dialectal irregularities. This amounts to explaining the difference between Mycenaean Greek and Proto-Indo-European with dialectal irregularities, because Latin, Old Indian or Hittite do not fit in the mould of Mycenaean Greek. On the whole we tend to disagree with most of Greenberg's conclusions, but we think the first chapter of *Language in the Americas* (1987) is always worth reading ever and ever again. It states (1987:16): “An eighth option is to explain irregular correspondences by positing dialectal mixture.” Greenberg lists ten ways how not to work with sound correspondences as the regular and recommendable method. He lists them in order to criticize the legitimacy of the comparative method. For our part, we read this more as a handbook to detect defective methodology.

What is more troublesome when it comes to the accuracy of PPM is that it also fails to account for some non-Min dialects, such as WenZhou. Here, with non-Min dialects, the problems seem to be weaker as regards their nature and depth but the relationship between the reconstruction and the dialectal data is still riddled with minor oddities. In other words, the reconstructions of PPM made by Baxter, Sagart and others (improperly called *Old Chinese* by them) account for a perimeter of Chinese dialects, that definitely excludes Min dialects and that includes a set of non-Min dialects that remains to be determined. This is the - skewed - basis on which external connections of Chinese can be made at the present time.

## 3. *Chinese and Austronesian*

For about twenty years, Sagart has been working on the hypothesis that Chinese is a close relative of Austronesian (AN). The most update arguments for this hypothesis are listed in his paper (2005) *Sino-Tibetan-Austronesian: An updated and improved argument*. For the sake of simplicity, we shall deal mainly with Chinese data.

Given the current state of the reconstruction of “Old Chinese”, such a perimeter as Sino-Tibetan-Austronesian is like trying to compare Mycenaean Greek with Uralo-Altaic, without knowing what Proto-Indo-European looks like and not knowing what the perimeter of Altaic is.

Among the potentially inherited morphological arguments, shared by Chinese and Austronesian, an interesting one is the “distributed action/intensive” morpheme. Conspicuously, Chinese examples are all from the Min dialect of FuZhou. This raises several questions. Apart from the obvious one, which is the level of intrinsic relevance of the examples, another one is to determine whether they are to be considered as a proof that Chinese, as a whole, is related to AN, a proof that FuZhou alone is related to AN, a proof that some or all Min dialects are related to AN, or a proof that some or all Chinese dialects have a substrate or adstrate of AN origin.

In other words, the issue is to know whether FuZhou is a Chinese dialect with a potential AN substrate/adstrate or whether it is an originally non-Chinese AN language with a huge Chinese cultural and lexical superstrate. This issue is not even hinted at in the paper. The monophylogenicity of Chinese dialects is a premise, which is widely accepted as “obvious” but which is currently *unproved*. Considering the problems entailed by Min dialects within the framework of present-day sinology, the monophyletic hypothesis cannot be taken as a given in our humble opinion

#### 4. The reconstruction of “Old Chinese” in Sagart (2002)

The data presented by Sagart in Canberra (2002) contain about seventy “Proto-Chinese” items. As we have previously underlined, the comparison of these “reconstructions” with Chinese dialects reveals major problems. Three present-day dialects are compared: Mandarin (BeiJing), ChaoZhou (a Min dialect), WenZhou (a non-Min dialect). At least in theory, WenZhou would be expected to compare more regularly to Mandarin (BeiJing) than ChaoZhou, but in practice, the situation is uneasy.

To begin with, if we look at word-final *\*-aʔ* in Sagart's “reconstruction”:

- 女 B10 ‘woman’: Sagart (2002) *\*nraʔ* ~ BeiJing nü3 ~ WenZhou nü22 ~ ChaoZhou n̄iŋ21
- 土 B17 ‘earth’: Sagart (2002) *\*thaʔ* ~ BeiJing tu3 ~ WenZhou: thoü21 ~ ChaoZhou thou21
- 鹵 B19 ‘salt’: Sagart (2002) *\*raʔ* ~ BeiJing lu3 ~ WenZhou l̄eu22 ~ ChaoZhou lou21
- 吐 B48 ‘vomit, spit’: Sagart (2002) *\*thaʔ* ~ BeiJing tu3 ~ WenZhou thoü21 ~ ChaoZhou thou31

All these words are reconstructed by Sagart's OC as rhyming with *\*-aʔ*. One would expect that, if these words with a quite simple structure #Caʔ were rhyming in OC, then they would most probably still rhyme in present-day dialects.

As regards tone, the final glottal stop accounts for BJ having tone 3. But, when we include other dialects, what we have is *three* different correspondences: 3 = 22 = 21, 3 = 21 = 21, 3 = 21 = 31. How should this split into three correspondences be explained? As regards the vocalic nucleus, what we have is *three* different correspondences: ü = ü = i, u = oü = ou, u = ëu = ou. How should this split into three correspondences be explained? In other words, whatever it really was sounding like, OC was certainly not like Sagart's “reconstruction”. This reconstruction is obviously unable to account for the phonetic reality of Chinese dialects. Moreover, the word ‘snake’ supposedly ending with *\*-a* without the final glottal stop has yet another vocalic correspondence:

- 蛇 B14 ‘snake’: Sagart (2002) *\*mla* ~ BeiJing she2 ~ WenZhou zei12 ~ ChaoZhou cua12

If we look at a cultural wanderwort like ‘tea’, originally a word meaning ‘leaf’, the comparative lexical evidence within Sino-Tibetan tends to show that the vowel was most probably *\*a* and we can see that this word displays the following correspondence:

- 茶 ‘tea’: Baxter *\*La* ~ BeiJing cha2 ~ WenZhou zo12 ~ ChaoZhou te12

This same regular correspondence OC *\*a* = a = o = e is shown in another word:

- 馬 ‘horse’: Baxter *\*mraʔ* ~ BeiJing ma3 ~ WenZhou mo21 ~ ChaoZhou be21

These regular items (tea, horse) show that *none* of the Chinese words (woman, earth, salt, vomit, snake) listed above which Sagart compares with Austronesian words rhyming with *\*-a* can possibly have had *\*-a* as main vowel. None of these words alleged as cognates can in fact be cognates at all. These words did not and do not even rhyme together.

The same problems appear with word-final *\*-uʔ* as reconstructed by Sagart:

- 腦 B3 ‘brain’: Sagart 2002 *\*nuʔ* ~ BeiJing nao3 ~ WenZhou nē22 ~ ChaoZhou nao21
- 水 B3 ‘water’: Sagart 2002 *\*hluʔ* ~ BeiJing shui3 ~ WenZhou si21 ~ ChaoZhou cui21

How can these words have ever rhymed in any Chinese dialect? Sagart’s “reconstruction” is not only inconsistent with Chinese dialects but it can be proved to be doubtless inadequate.

A possible conclusion is that this reconstruction has been made to sound like Austronesian but the lexical data of Chinese dialects do not fit in that procustean mould. Whatever OC may have been and whatever the relationship between Chinese, Sino-Tibetan and Austronesian might be, the examples alleged by Sagart as evidence of genetic relationships are in fact indications that Chinese contains a huge bulk of vocabulary that has *no clear genetic connections* with Austronesian.

## 5. Conclusion

An interesting historiography of Sino-Tibetan was written in 2005 by George van Driem. As the paper shows, Sino-Tibetan is mainly a negative and residual entity, which includes all the languages which had been originally lumped together by Klaproth in 1823 on account of their being neither indo-european nor hamito-semitic and which have never been proved considered to be unrelated to Chinese or Tibetan. Sino-Tibetan is not a positive entity. It still includes after two centuries of gradual reduction those languages of which it has never been thought they should be handled with better results within the framework of another family. As van Driem states, “Although a reconstruction of Proto-Kiranti, for example, is available (Opgenort 2005), no reconstructions are available for most branches”. Van Driem is obviously uncomfortable with Sagart’s claim that Austronesian is a legitimate relative of Chinese. He accepts the idea of an “ancient contact situation”. The major problem facing van Driem is that Sino-Tibetan has no positive groundwork and for that matter it is impossible to show that Austronesian does *\*not\** belong there.

In fact, the hypothesis proposed by Sagart is an unpinned grenade thrown in a castle of cards. What holds the castle of cards together are mainly typological features such as frequent tonogenesis and the so-called monosyllabic pattern. It is well known that typological features are not necessarily equivalent to inherited features and that the monosyllabicity of many modern far-eastern Asiatic languages is the result of phonetic attrition.

The current state of Sino-Tibetan make it neither possible nor impossible to assess the position of Austronesian vis-à-vis that family. It is fascinating to see that Sagart rejects Matisoff’s proto-Sino-Tibetan because it is neither explicit, nor reliable nor falsifiable but he nevertheless undertakes to prove that Austronesian is a legitimate member of that unproved super-family. Sagart’s theory makes it necessary to understand what corresponds to cognacy and borrowing in Sino-Tibetan, something that has never been done so far. It is unclear what will remain of the castle of cards when the grenade has finished to send shock waves. Chinese may even not survive the blow-up. If Austronesian once had relatives on the Chinese mainland, then something must be recoverable in present-day “Chinese” dialects. The issue is ultimately to understand what Min dialects really are. Are they genetically Chinese or genetically Austronesian? Or something else? This is also what is at stake with the “Sino-Austronesian” hypothesis from the point of view of Chinese.

*References*

Baxter, William. 1992. *A Handbook of Old Chinese Phonology*. Berlin and New York: Mouton de Gruyter.

Greenberg, Joseph Harold. 1987. *Language in the Americas*. Stanford University Press.

Sagart, Laurent. 1999. *The Roots of Old Chinese*. John Benjamins: Amsterdam and Philadelphia.

Sagart, Laurent. 2005 (first presented in 2002). Sino-Tibetan-Austronesian: An updated and improved argument. <http://halshs.archives-ouvertes.fr/docs/00/08/50/59/PDF/canberra.pdf>

Van Driem, George. 2005. *Sino-Austronesian versus Sino-Caucasian*.  
[www.eastling.org/paper/Driem.pdf](http://www.eastling.org/paper/Driem.pdf)