**Answer key to Proto-Chinese reconstruction exercise**

In some ways, the divergence of a common spoken Chinese language into the different dialects now that are mostly mutually unintelligible is comparable to the development of the Romance languages from Vulgar Latin. Now let’s see a few examples. Ignore tones for now.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Guangzhou** | **Taiyuan** | **Suzhou** | **Wenzhou** | **Beijing** | **Meaning** |
| sam | sæ̃ | sE | sa | san | **三three** |
| tam | tæ̃ | tE | ta | dan | **担 carry** |
| lam | læ̃ | lE | la | lan | **藍blue** |
| fan | fæ̃ | vE | va | fan | **飯meal** |
| tsan | tsæ̃ | tsE | tsa | zan | **贊 praise** |
| tan | tæ̃ | tE | ta | dan | **單 alone** |

1 What are the correspondences? Can you list all of them?

* The easier correspondences and their reconstructed forms are:

s-s-s-s-s \*s

t-t-t-t-t \*t

l-l-l-l-l \*l

ts-ts-ts-ts-ts \*ts

But note that it is also possible that the original form changed in all these daughter languages, and we don’t have enough data to show this possibility here.

* Another correspondence of vowels is a-æ̃-E-a-a. Since the majority languages have a, and æ̃ is also quite similar to a (cf the a in “father” and the a in “fathom”), we can reconstruct the original form as \*a for now. Unless there is enough data to show otherwise.
* Also we have f-f-v-v-f. So it is \*f or \*v. We do not have enough evidence here. But if you look at Grimm’s Law, proto-IE \*b, \*d, \*g became p, t, k in Germanic. Thus it is possible that \*v>f in the same way. This is called devoicing.

Now let’s look at the last two correspondences:

* The last three words have n-ф- ф- ф-n. Since it is very unlikely that the sound n developed out of nothing, then the original reconstructed sounds should be \*n
* The first three words: m-ф- ф- ф-n. Similarly, the more likely original forms would be either \*m or \*n.

2 What are the reconstruction possibilities?

* Suppose it is \*n, then the original words would look like those in Beijing (except for fan, which would be \*van). Then somehow the first three words in Cantonese changed to having –m as the ending, while the remaining three words didn’t change. This is called a *split* of sound. A split is always *conditioned*, i.e. there would be a condition for the split. For example, if you recall the proto-Romance k split into ch and k in French. The condition is the vowel that follows the original k. In our sample words, if it is a vowel like “e”, then \*k>ch (this is called palatalization). Otherwise \*k>k.

Therefore if we posit \*n in this Proto-Chinese example, then we must find a condition for \*n>m. If we look at the vowel in each word, they are all the same. *No condition whatsoever can be found.*

Therefore this hypothesis of an original \*n is not very likely.

* Therefore we are left with the only choice, i.e. \*m. Then the original words would look like those in Cantonese (except for fan, which would be \*van). Then \*m>n in Beijing. Since there is always a –n of a different source, we say that syllable-ending \*m and \*n in Beijing merged into n.

*Different from splits, a merge of sounds does not need any condition*, although there often are some conditions.

* Therefore the reconstructed words are \*sam, \*tam, \*lam, \*van, \*tsan, \*tan.

A historical sketch of the development from this proto-language into the present languages is:

* In these few words, nothing except \*van>fan changed in Cantonese. (NOTE that we are only using a tiny fragment of the whole language. Thus we cannot say that Cantonese didn’t change. It is incorrect to say that.)
* In Beijing, \*m and \*n merged. \*van>fan. Nothing else changed. The consequence of the merge resulted in a pair of homophonous words which were formerly different, i.e. \*tam and \*tan.
* In Taiyuan, probably like in Beijing at an earlier stage, \*m>n, and then n will be added to the vowel, instead of being just an independent sound. Also the vowel changed slightly as well \*a>æ. We can imagine that the n is in the process of being lost.
* If either \*m>n and n is lost or \*m and \*n are lost, we get a language like Wenzhou.
* If another change of the vowel \*a>E took place, then we get a language like Suzhou.