On a mode of borrowing from Middle Chinese into Proto Tibetan

(A new look at the problem of the relationship between Chinese and Tibetan)

FERLUS, Michel

Abbreviations:

MC  Middle Chinese (Karlgren Ancient Chinese): the stage of the Qie$^4$ Yun$^4$
     reflected in the Yun$^4$ Jing$^4$.
OC  Old Chinese (Karlgren Archaic Chinese): the stage of the rhymes of the Shi$^1$
     Jing$^1$.
WT  Written Tibetan
PT  Proto Tibetan (The stage just before the Middle Chinese influence)
PST Proto Sino-Tibetan (in a restricted sense)
TB  Tibeto-Burman

Introduction:

It is well known that Chinese and Tibetan are genetically related. But it has been
asserted by some scholars that an important part of the supposed common vocabulary
represents an ancient layer of borrowings from Middle Chinese into Ancient Tibetan.
Depending of the point of view, Chinese and Tibetan can be considered as more or less
genetically related. It is important to separate the loanwords from inherited vocabulary in
order to evaluate the closeness of the relationship.

The study of loanwords depends on the relationship of the languages in contact.
When the languages are not genetically related, it is in general relatively simple to sort out
loanwords. In this process the borrowing concerns the whole of one word. But when
genetically related languages are in contact with a degree of intercomprehension and in a
hierarchical relationship of prestige, more complex types of influence can occur. This kind
of process works only with one segment in a word, the borrowing affects one syllable in a
disyllabic word or one constituent (consonant, vowel or rhyme) in a syllable. This
phenomenon has been observed, with more or less degree of importance, by the author in
several cases of linguistic contacts in the Southeast Asian area.

This process will be illustrated here by borrowings, or partial borrowings, from
Middle Chinese (MC) into Proto Tibetan (PT), a stage of the language that preserved a
part of its vocabulary still relatively closed to Proto Sino-Tibetan (PST). Results of these influences are reflected in Written Tibetan (WT).

Reminder of the theory of monosyllabization from OC to MC

Before any further developments it seems useful to remind our theory on the phenomenon of monosyllabization that happened between OC and MC (Ferlus 1998). Without the knowledge of it, the following explanations will not be clearly understand.

Old Chinese was a disyllabic language. It means that one part of the vocabulary was constituted of monosyllabic words, while the other part was constituted of disyllabic words, more precisely of the sesquisyllabic type (according to J. Matisoff’s definition). This type is still widely represented in many austroasiatic languages of Southeast Asia (Ferlus 1996). A sesquisyllable is a type of disyllable made of a main syllable preceded by a presyllable. The main syllable is similar to a monosyllabic word while the presyllable is a reduced and unstressed syllable without phonemic vowel. The presyllable can be a morphological prefix as well as a neuter element without any signification.

monosyllable: CV(C)
sesquisyllable: C-CV(C)

Sesquisyllables of OC developed a tenseness (T) while by contrast monosyllables developed a laxness (L). Then, sesquisyllabic words shifted in monosyllabic by loosing presyllables. Consequently, the former contrast of syllabic type C-CV(C) vs CV(C) was replaced by the new contrast tense vs lax (T/L). This phenomenon was associated to a vowel splitting showing vowel lowering in T syllables and vowel rising in L syllables. Later, at a second step, after these changes, the softening of medial -r- came blurring the situation. It is the stage of MC characterized by the famous four divisions system: the T syllables belong to division I/IV (no medial -r- in OC) or to division II (medial -r- in OC), while the L syllables belong to division III (with or without medial -r- in OC) characterized by the famous yod of Karlgren’s reconstructions (1957).

<table>
<thead>
<tr>
<th>Old Chinese (OC)</th>
<th>Middle Chinese (MC)</th>
<th>divisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-CV(C) (tenseness)</td>
<td>CV(C) / T (vowel lowering)</td>
<td>I/IV (−r) or II (+ r)</td>
</tr>
<tr>
<td>CV(C) (laxness)</td>
<td>CV(C) / L (vowel rising)</td>
<td>III (± r)</td>
</tr>
</tbody>
</table>

In my system the symbol [³], called here the proteus, is the mark of the division III. It indicates a certain lowering and centralization of the vowel associated with (what I suspect to be) a breathy voice. The symbol [⁺], called here the spirans, is the mark of the division II. It results from the softening of OC medial -r- and seems to indicate a kind of spirantized velar sound. The Division II is nothing more than a variety of division I (and IV). In L syllables the softening of OC medial -r- has been absorbed by the breathiness of the vowel, it is the reason why it don’t exist a variety of division III, as division II is for division I. The division IV is in complementary distribution with the division I and seems to be just a device to describe the single MC front diphtong. There is no mark to indicate the divisions I/IV.
Schema résumé:

<table>
<thead>
<tr>
<th>tense syllables</th>
<th>no medial -r-</th>
<th>medial -r-</th>
</tr>
</thead>
<tbody>
<tr>
<td>(OC)</td>
<td>MC</td>
<td>(OC)</td>
</tr>
<tr>
<td>( - )</td>
<td>I/IV / T</td>
<td>( -r- )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>lax syllables</th>
<th>lax</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-j- )</td>
<td>III / L</td>
</tr>
</tbody>
</table>

Some examples: The presentation is taken from Baxter (1992), I only added my own phonetic interpretation between square brackets.

T I 納 na<sup>3</sup> < nop [nʌp] < *nup [T(C-)nup] "send in" (695h)
L III 入 ru<sup>3</sup> < nyip [nʌjip] < *n-j-up [L-nup] "enter" (695a)
T IV 銘 ming<sup>2</sup> < meng [mɛn] < *meng [T(C-)men] "inscription" (826d)
L III (¬IV) 名 ming<sup>2</sup> < mjieng [m³jen] < *m-j-eng [L-men] "name" (826a)
T I 投 gu<sup>3</sup> < ku<sup>x</sup> [kɔʔ] < *kaʔ [T(C-)kaʔ] "thigh" (51a)
Tr II 假 jia<sup>3</sup> < kæ<sup>x</sup> [kʰəʔ] < *k-r-aʔ [T(C-)kraʔ] "false, simulate" (33c)
Lr III 蓑 ju<sup>3</sup> < k jo<sup>x</sup> [kʰjæʔ] < *k-rj-aʔ [L kraʔ] "round basket" (76j)

Before this complete and structural phenomenon of monosyllabization that affected the whole sesquisyllabic vocabulary, it could happen a kind of slow and aleatory monosyllabization that affected the vocabulary word by word.

Brief presentation of the mode of borrowing from MC into Proto Tibetan

Proto Tibetan (PT) is defined as the stage of the Tibetan language just before the MC influence. PT could be as well called pre Old Tibetan.

One language A (here MC) is in a dominating position with a genetically related language B (here PT). Language A is regarded as prestigious by speakers of B who, by a kind of affection, are led to imitate some characteristic sounds of A unknown in B. This results in a phonetic compromise, a segment of an A word being borrowed and put in place of the corresponding segment of the cognate B word. The characteristic sounds of MC that don't exist in PT are the segments (rhymes or main syllables) involved by the division III (presumed breathiness marked by the proteus [ʰ]) and the division II (spirantized velar sound marked by the spirans [ʕ]). Speakers of PT tried, unconsciously or not, to imitate these unfamiliar sounds of MC because their pronunciation by Chinese speakers was felt as more prestigious. But, at the difference of the usual process of borrowing that concerns the entire word, ancient speakers of PT borrowed only the segment bearing the characteristic sound in a MC word.

This process will be first illustrated with the numbers "one" to "ten" and "hundred".

Comparison of Tibetan and Chinese numerals


The Chart 2 is the reference chart for the following demonstrations.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ar C / An C</td>
<td>PST</td>
<td>OC &gt; MC</td>
<td>TB</td>
<td>TB</td>
<td>EMC</td>
<td>OC &gt; MC</td>
<td>OC &gt; MC</td>
</tr>
<tr>
<td>1 zhi¹ 赤</td>
<td>gtyik</td>
<td>tijk &gt; tʃjæk</td>
<td>g-tyik</td>
<td>t(y)ik</td>
<td>tʃajk</td>
<td>(tʃek &gt; tʃenk)</td>
<td>gcig</td>
</tr>
<tr>
<td>2 a³ 二</td>
<td>gnyis</td>
<td>njidh &gt; nži</td>
<td>g-nisy</td>
<td>nìh</td>
<td>nìts &gt; n̄ᵣ RequestContext</td>
<td>b̄ni[j]-s &gt; n̄iy RequestContext</td>
<td>gnyis</td>
</tr>
<tr>
<td>3 san¹ 三</td>
<td>gsum</td>
<td>sám / sâm</td>
<td>g-sum</td>
<td>sám</td>
<td>sum &gt; [sám]</td>
<td>ñksum &gt; [sám]</td>
<td>ās-hlim &gt; sam</td>
</tr>
<tr>
<td>4 s¹ 四</td>
<td>blyid</td>
<td>sjidh &gt; sì</td>
<td>blyiy</td>
<td>sìh</td>
<td>s(p)jj[s] &gt; sì̄R RequestContext</td>
<td>b̄s-hli[j]-s &gt; sījH RequestContext</td>
<td>bzhí</td>
</tr>
<tr>
<td>5 wu³ 五</td>
<td>ḷngay</td>
<td>ngagR &gt; ngwo</td>
<td>l-nga</td>
<td>l-qa</td>
<td>nga&gt; ngu</td>
<td>ñnja&gt; n̄a透露</td>
<td>āŋa &gt; nguX透露</td>
</tr>
<tr>
<td>6 liu⁴ 六</td>
<td>ḷjök / ljuk</td>
<td>djakw &gt; ljk</td>
<td>d-ruk</td>
<td>d-ruk</td>
<td>luwk</td>
<td>C-rjuk &gt; ljuwk</td>
<td>ḷruk &gt; luwk透露</td>
</tr>
<tr>
<td>7 q¹ 七</td>
<td>shnjis</td>
<td>tshjit &gt; tʃjet</td>
<td>s-nis</td>
<td>s-nis</td>
<td>tʃit</td>
<td>tʃs jit &gt; tʃshit透露</td>
<td>ḷtsit &gt; tsit透露</td>
</tr>
<tr>
<td>8 ba¹ 八</td>
<td>prwat / pwât</td>
<td>priat</td>
<td>b-r-gyat</td>
<td>poit/pet</td>
<td>(pret &gt; pet)</td>
<td>p̄pret &gt; p̄pet透露</td>
<td>āpr[e][e]t &gt; peat透露</td>
</tr>
<tr>
<td>9 jiú³ 九</td>
<td>dkwajyw</td>
<td>kjagw &gt; kjau</td>
<td>d-kuw</td>
<td>kuw</td>
<td>k̄wju&gt; kjuw透露</td>
<td>ḷk̄wu &gt; k̄uw透露</td>
<td>b̄ku &gt; kjuw透露</td>
</tr>
<tr>
<td>10 shi² 十</td>
<td>djəp / žəp</td>
<td>gip</td>
<td>gip</td>
<td>gip</td>
<td>gip</td>
<td>gip &gt; dzjip透露</td>
<td>ḷgip &gt; dzjip透露</td>
</tr>
<tr>
<td>100 bai³ 百</td>
<td>p̄ak / pok</td>
<td>pria’</td>
<td>prak &gt; p̄uk</td>
<td>prya</td>
<td>r-gya</td>
<td>paikj/p_ejk透露</td>
<td>prak &gt; p̄ek透露</td>
</tr>
</tbody>
</table>
## CHART 2

<table>
<thead>
<tr>
<th>OC &gt; MC</th>
<th>PT + MC</th>
<th>(*_{intermediate} form)</th>
<th>WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  zhi(^1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ltek &gt; tɕ(^a)ek</td>
<td>k[tek] + tɕ(^a)ek &gt; *ktɕ(^a)ek</td>
<td>gcig</td>
<td></td>
</tr>
<tr>
<td>2  er(^4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lnits &gt; n(^a)i(^j)h</td>
<td>k[nits] + n(^a)i(^j)h &gt; *kn(^a)i(^j)h</td>
<td>gnyis</td>
<td></td>
</tr>
<tr>
<td>3  san(^1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tksam &gt; [sam]</td>
<td>ksum &gt;</td>
<td>gsum</td>
<td></td>
</tr>
<tr>
<td>4  si(^4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lslits &gt; s(^a)i(^j)h</td>
<td>p[sits] + s(^a)i(^j)h &gt; *ps(^a)i(^j)h</td>
<td>bžhi</td>
<td></td>
</tr>
<tr>
<td>5  wu(^3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tlna &gt; gŋ(^a)</td>
<td>lŋa &gt;</td>
<td>lŋa</td>
<td></td>
</tr>
<tr>
<td>6  liu(^4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lruk &gt; l(^a)u(^w)k</td>
<td>truk &gt;</td>
<td>drug</td>
<td></td>
</tr>
<tr>
<td>7  qi(^1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ltshit &gt; tsh(^a)lt</td>
<td></td>
<td>(bdun)</td>
<td></td>
</tr>
<tr>
<td>8  ba(^1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tpret &gt; p(^a)c(^t)</td>
<td>pr[et + p](^a)c(^t) &gt; *pr(^a)c(^t)</td>
<td>brgyad</td>
<td></td>
</tr>
<tr>
<td>9  jiu(^3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lkwu &gt; k(^a)u(^w)</td>
<td>t[ku] + k(^a)u(^w) &gt; *tk(^a)u(^w)</td>
<td>dgu</td>
<td></td>
</tr>
<tr>
<td></td>
<td>or tku &gt;</td>
<td>dgu</td>
<td></td>
</tr>
<tr>
<td>10 shi(^2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lgip &gt; dz(^a)ip</td>
<td>p[gip] + dz(^a)ip &gt; *dz(^a)i(p)</td>
<td>bcu</td>
<td></td>
</tr>
<tr>
<td>100 bai(^3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tprak &gt; p(^a)æk</td>
<td>pr[ak + p](^a)æk &gt; *pr(^a)æk(k)</td>
<td>brgya</td>
<td></td>
</tr>
</tbody>
</table>

- "One ~ alone":
  PT "one" *ktek, WT gcig.
  OC "alone" *tjek [l-tek] > MC tseyk [tɕ\(^a\)ek] > zhi\(^1\) 鼻 (1260c), not attested in Baxter (1992). The actual word for "one" is yi\(^1\) ㄧ < MC ʔjit [ʔ\(^a\)it] < OC ʔjit [l-zit].
  PST *ktek. The presyllable k-, reconstructed on the basis of WT g-, was lost in pre-OC times by aleatory monosyllabization. The division III of MC led to reconstruct a monosyllable in OC.

  During the interferences of MC forms with PT forms, the main syllable in PT k-tek was replaced by the unfamiliar pronunciation for Tibetan speakers of MC tɕ\(^a\)ek. The combination k[tek] + tɕ\(^a\)ek rose to an hypothetic intermediate form *ktɕ\(^a\)ek well represented by WT gcig.

- "Two":
  PT *knits, WT gnyis.
  OC *njits [l-njits] > MC nyij\(^h\) [n\(^a\)i\(^j\)h] > er\(^4\) 二 (564a), not attested in Baxter.
  PST *knits. The presyllable k-, reconstructed on the basis of WT g-, was lost in pre-OC times by aleatory monosyllabization, the division III of MC led to reconstruct a
monosyllable in OC. The final -ts changed into -js > -jḥ by final cluster simplification (Baxter 1992: 568-9).

The main syllable in PT knits was replaced by the unfamiliar pronunciation of MC ə>jḥ. The combination k[nits] + ə>jḥ (with the possibility of a pre-MC form ə>is) rose to an hypothetical intermediate form *knə>jḥ well represented by WT gnyis.

- "Three":
  PT *ksum, WT gsum.
  PST *ksum.

The pronunciations of MC sam as well as any other MC forms in -ʌm (see below), all belonging to the divisions I or II, were not exotic for Tibetan speakers. So they did not need to imitate it and the WT gsum derives directly from PT *ksum.

Discussion about "three": The problem risen by the reconstruction of "three" and its word family is a very complex one. It has been highly treated by Sagart (1999: 148-152). To discuss the ideas of the author will lead us too far from the present subject. I will just give here briefly my point of view.

L. Sagart proposed two forms for "three": OC *as-hlîm > MC sam > sanl  (... for the simple graph and OC *as-hlîm > MC tshom > canl 参 for the complex graph or da4 xie3. First of all, I consider that the rhymes -um / -up must be reconstructed, the changes -um > -im and -up > -ip occurred after OC times. Aside the basic form OC Tksum > sanl  (... "three", the word family comprises the MC meaning of 臺 (read canl), MC tshom [tsʰʌm] "three horses in a team", and the both MC meanings of 参 (read canl), MC tshom [tsʰʌm] "three, a triad" and (read shenl) MC srim [sʰim] "the triad star of Orion". The character 参 (read sanl) is also used until to day as complex graph for "three". It must be noticed that MC rhymes in tshom [tsʰʌm] (division I) and in srim [sʰim] (division III) are regular in respect to OC rhyme -um [-um]. For these two words I propose the reconstructions OC *srunt [Tksum] > MC tshom [tsʰʌm] and OC *srunt [Lsrunt] > MC sim [sʰim] that I consider as secondary forms of OC Tksum. In résumé:

pre-OC ksum > OC Tksum > MC sam (irr.) > sanl (... "three".
pre-OC krsum > (metathesis of -r-) OC Tksum > MC tsʰʌm 参 "three, a triad" > canl (reading by the way of the meaning "take part, visit"), also MC tsʰʌm 臺 "three horses in a team".
pre-OC krsum > (loss of k- and metathesis of -r-) OC Lsrunt 参 "the triad star of Orion" > MC sʰim > shenl (reading by the way of the meaning "ginseng").

The change of pre-OC krsum into OC Tksum or Lsrunt by metathesis of -r- from the presyllable to the main syllable is, of course, purely hypothetic. But this phenomenon of metathesis could help us to understand the curious fugacity and the intrusive behaviour of some OC medial -r- and the fact that items with or without this medial can occur in the same phonetic serie. Nevertheless the idea of an ancient metathesis of -r- is supported by some lexical correspondences between WT and OC. Let us compare WT rdul "dust" with chen2 鍕 < MC drin [qʰin] < OC *drin [drʰin] (< pre-OC dril) "id." (example seen in Coblin 1986: 68).
I propose that the archaic character for 参 had previously the meaning "three horses in a team" because this notion was more familiar to peoples than those of "triad of Orion". The meaning of three horses is expressed by the upper part of the character that indicates rather clearly the three horse's heads, contra some other scholars who prefer to see the three stars in it (why the three stars would be tied on ?). The lower part of the character has been sometime interpreted as the phonetic element, but neither the element zhen\(^3\) 参 (OC rhyme -in) nor shan\(^1\) 参 (OC rhyme -am) fits phonetically with can\(^1\) 参 (OC rhyme -um). For myself, I prefer to see in the lower part of the archaic character for 参 the image of reins hung with ornaments.

The ancient pronouciation of 参 was used to derive numerous other characters that belong to the phonophoric serie GSR 647.

- "Four":
  PT *psits, WT bzhi.
  OC *s(p)jjj/1ts [\(^1\)-slits] > MC sij\(^H\) [s\(^9\)i\(^j\)h] > si\(^d\) 四 (518a).
  PST *plsits ~ *pslits. The presyllable p- (may be a prefix ?) reconstructed on the basis of WT b- was lost in pre-OC times. The medial -l- is justified by occurrences in some Tibeto-Burman languages.

  The main syllable in PT psits was replaced by a corrupted form of the unfamiliar pronunciation of MC s\(^9\)i\(^j\)h. The combination p[sts] + s\(^9\)i\(^j\)h rose to an hypothetic intermediate form *ps\(^9\)i\(^j\)h rather well represented by WT bzhi.

- "Five":
  PT *l\(\eta\)a, WT inga.
  OC *nga? [T\(\eta\)a?] > MC ngu\(^X\) [\(\eta\)p?] > wu\(^3\) 亻 (58a).
  PST *l\(\eta\)a.

  As for "three", WT inga derives directly from PT without interference with MC.

- "Six":
  PT *truk, WT drug.
  OC *C-rjuk [\(^1\)-ruk] > MC ljuwk [1\(^a\)uwk] > lur\(^4\) 六 (1032a).
  PST *truk ~ t-ruk. The first element t- was lost during pre-OC times, it means that tr- must have been a kind of disjoined cluster. In proto Thai proper "six" is reconstructed as *hruk, the voiceless being an irrefutable proof of an ancient presyllabic element.

  As for "three" and "five", the WT form derives directly from PT.

- "Seven":
  PT *pdun, WT bdun.
  OC *thsjit [\(^1\)-ts\(^h\)it] > MC tshit [ts\(^h\)it] > qi\(^l\) 七 (400a).

  Pre-OC (for PST) *snit ~ s\(^h\)nit. One can remarks that the rhyme -it(s) exits in knits "two", plsits ~ pslits "four (2+2)" and snit ~ s\(^h\)nit "seven (5+2)".

  The Tibetan and Chinese forms are not genetically related.

- "Eight":
  PT *pret, WT brgyad.
  OC *pret [\(^1\)pret] > MC pet [p\(^z\)et] > ba\(^l\) 八 (281a).
PST *pret.

The rhyme -et in PT pret was replaced by the unfamiliar pronunciation of the segment -et of MC pśet. The combination pr[et + p]ś et rose to an hypothetical intermediate form *pret rather well represented by WT bryad. The segment -gad is the result of the interpretation of -et in the phonetic system of Tibetan. Some scholars have interpreted -g as an epenthetic element, but it must be remarked that epenthetic sound appears usually at the juncture of two syllables and it was not the ground in the Tibetan word for "eight".

- "Nine":
  PT *tku, WT dgu.
  OC *kuwu? [k'wu? ~ kuw?] > MC kjuwX kauw > jiu3 九 (992a). The reconstruction of a labiovelar before a high rounded vowel by Baxter is surprising in spite of a firm demonstration, so I will propose an alternate form k'wu?.
  PST *tku? ~ *tku.

  At first sight, WT dgu derives directly from PT tku but the voiced velar -g- does not fit perfectly with PST and could be a result of the laxness on MC form. If so, the combination t[ku] + k'auw rose to an hypothetical intermediate form *tk'auw represented by WT dgu.

- "Ten":
  PT *pgip (?), WT bcu.
  PST *pgip. The presyllable p- (may be a prefix ?) is reconstructed on the basis of WT b- assuming that these forms are related. I suppose that the combination p[gip] + dzəip rose to an hypothetical intermediate form *pdzəip. Does the rhyme WT -cu could represent MC dzəip after the loss of final -p? I must confess that I am not absolutely sure of that.

- "Hundred":
  PT *prak, WT brya.
  PST *prak.

  The demonstration about "hundred" is parallel to those of "eight". The combination pr[ak + p]əak rose to an hypothetical intermediate form *prəak > prə rather well represented by WT brya. The loss of final -k is unexplained but not isolated.

Conclusions

Among the eleven correspondances between Tibetan and Chinese about numerals, nine can be considered as good one. Out of that, occurrences for "seven" are not cognates and correspondence for "ten" is not absolutely sure.

In correspondances about "three", "five" and "six", the WT forms derive directly from PT without MC interference. They must be considered as pure inherited correspondances.

In correspondances about "one", "two", "four", "nine" (presumably) and "ten" (if related) the MC monosyllable replaced the main syllable in PT form, but as for "eight"
and "hundred" the segment replaced was the rhyme. They are what can be called
corrupted (or modified) inherited correspondances in which the Tibetan word is the result
of a compromise between an inherited form and an acquired segment while the Chinese
term remains unchanged. This kind of correspondances can be placed between pure
inherited correspondances and acquired correspondances by borrowings.

This mode of borrowing working only with a segment of the word (main syllable or
rhyme) is the consequence of a particular situation. The two languages in contact are
genetically related with a certain degree of intercomprehension and in a hierarchical
relation of prestige. The Chinese language of MC times being in a dominating position
was regarded as prestigious by speakers of Tibetan who were led to imitate, by a kind of
affectation, some characteristic segments (sounds of division II and III) unknown in
Tibetan.

Considerations have been limited to numerals that behave in general as a whole set,
but the analysis could be enlarged with profits to other domains of Tibetan vocabulary.

I propose to call hypercorrection by affected imitation this special process of
borrowing. Such layers of borrowings have never been clearly identified in historical
comparative studies.

Selected bibliography

Baxter, William H. 1995. "A stronger affinity...than could have been produced by
accident": a probabilistic comparison of Old Chinese and Tibeto-Burman. The
Chinese Linguistics monograph series no. 8.

Mouton de Gruyter.

Baxter, William H. 1993. Some phonological correspondances between Chinese and
Tibeto-Burman. 26th ICSTLL. Osaka, Japan.


seminar on East and Southeast Asian linguistics: 60-100. Tokyo.

Bodman, N.C. 1980. Proto-Chinese and Sino-Tibetan: data towards establishing the
nature of the relationship. Contributions to Historical Linguistics: Issues and

Coblin, Weldon South. 1986. A Sinologist's Handlist of Sino-Tibetan Lexical
Comparisions. Monumenta Serica Monograph Series XVIII. Nettetal: Steyler.
[Ouvrage de référence]

Ferlus, Michel. 1996. Evolution vers le monosyllabisme dans quelques langues de l'Asie
du Sud-Est. Conférence donnée à la Société de Linguistique de Paris, séance du 23
novembre 1996.

41


Michel Ferlus / CRL Asie Orientale, 54 blv Raspail, 75006 Paris, France / <ferlus@ehess.fr>