

JOSÉ ANDRÉS ALONSO DE LA FUENTE

Written Manchu *šumin* ‘deep’

Abstract

The main goal of this paper is to discuss the etymology of Written Manchu *šumin* ‘deep’ (§§1.1–1.3.1). This word is usually connected with Ewenki *suŋta* ‘deep’ (§2.1.1). Benzing and Doerfer proposed partial solutions to account for the obvious irregularities existing between one and another form (for starters, irregular sound correspondences *š-* vs. *s-*, *-m-* vs. *-ŋt-*, etc.). Though reasonable and to some degree plausible, such a proposals still need some refinement (§2.1.2). Additional explanations shall be provided with regard to Jurchen data (§§1.3.2–1.3.3), general Tungusic derivational morphology (§2.1.3), and external comparisons (§§2.2.1–2.2.2).

Keywords: Manchu language, etymology, *šumin*, deep

1.1. Written Manchu [= WM] *šumin* is a very common word meaning ‘deep’. It is related to *šumji* ~ *šumci* ‘depth; sunken, submerged’, *šumikan* ‘rather deep’ and *šumila-* ‘to deepen, become deep’ (vid. i.a. HM 451b, CMEL 264b, MMC 447, PMRS 688a, TFAXB 239 s.vv. <siumin>, <siumila->) by regular derivative mechanisms involving the denominal nominal suffix *-ci* as in *fomo.n* ‘foot wrappings’ → *fomo+ci* ‘stockings’, the meliorative *-kAn* as in *hata.n* ‘strong’ → *hata+kan* ‘rather strong’, and the iterative *-lA* as in *nicu-* ‘to close one’s eyes’ → *nicu-la-* ‘to blink’, respectively. *Mittelsilbenschwund*, the process whereby an unstressed vowel in an open second syllable of a trisyllabic and longer word is systematically lost in Manchu (especially if the vowel is /i/ or /u/), applies to *šumji* ~ *šumci*, i.e. */šumĩ+ci/ > *šumci*, the variant *šumji* being the result of secondary voicing assimilation of /c/ to /m/, but neither to *šumila-* since this is a verbal formation, nor to *šumikan*, perhaps a recent derivative which was created when *Mittelsilbenschwund* was not active anymore.

1.2. Spoken Manchu and Sibe have an exact match for *šumin* id. (*vid. i.a.* Sanjiazhi dialect = MSM 120). Yamamoto's variant *sumin* with no initial palatal (CDSM 127 [2617]) bears witness to the merging of /š/ with /s/ (this is also common in Sibe, *vid. i.a.* Norman 1974: 162). In Muromski's Kulja Sibe materials we find the variant *šumiñ* (Kałużyński 1977: 255). A slight palatalization of /n/ in the surroundings of /i/ and other palatal sounds seems to be a common feature in these materials, *cfr. teñi* and *teni* 'just, then', *ixañ* and *ixan* 'cow, bull; cattle', *ejuñ* and *ejun* 'older sister', etc.

1.3.1. To the best of my knowledge, there is no cognate in Old Jurchen corresponding to WM *šumin*. As for Late Jurchen [= LJ], what seems to be two competing forms are attested: A 說迷 (šwemi´) †šomi [153], and H <侃劉受> 舒迷吉 (šymi´ki) †šumigi [695].¹ There are some questions to deal with before we can directly compare these forms with the corresponding Manchu-Sibe words mentioned above.

1.3.2. At least two additional instances in LJ reflect the same disparity A */o/ vs. H */u/ or WM /u/ regarding the stem vowel (the sound correspondence A */u/ vs. H */o/ is far more common, but equally hard to explain): A 莫答 (maw`ta´) †moda 'bend (of the river)' [240] vs. WM *mudan* id., A 好沙 (xaw`ša) †haoša 'paper' [563] vs. H <凡各> (= A) †hauša [222], *cfr.* WM *hōšan* id. and Sibe *hošin* ~ *ha'ušan* 'paper offerings used in ancestor worship ritual'. The latter reconstruction appears to be incorrect, if it is confronted with other instances containing the diphthongoid */au/: A 掃伏 (saw`fu´) †šaufu 'saddle-cushion' [619], *cfr.* WM *soforo* id., A 道力哈 (taw`li`xa) †dauriha 'to capture' [821], *cfr.* WM *duri-* 'to steal, snatch away from' or A 朝哈 (tš`hew`xa) †čauha 'army' [659], *cfr.* WM *cōha* id. Then, Kane's †haoša 'paper' should be emended to †hauša 'paper' (*pace* Jīn 1984: 153 who also reads */hao/, but in agreement with Kiyose's reconstruction in H, see rationale in 52-53). As for the former, namely †moda, it can be traced back to Proto-Tungusic [= PT] */mu(u)daⁿ/ id. (TMS I.542a-b). From this perspective, LJ */o/ is irregular. However, it also points out that the problem may lie on the side of the Chinese transliteration rather than be connected to Jurchen (historical) phonology, since the sound change chain PT */u/ > Jurchenic */o/ > Manchu /u/ is unheard of.

As for A †šomi, external evidence cannot help much here, as we shall see in the following section. The character used to transcribe LJ */šo/ turns out to be different from the one attested in the other words containing such a sequence: A 灼羅 (šew`lo or šo`lo) †šoro 'cage' [554], A 過灼 (kwəšew´ or kwəšo´) †gošo 'bitter' [231], *cfr.* WM *šoro* and *gosihon* id., and H <哈> 安朔 (anšwaw´) †amšo 'eleven' [646] (Jīn 1984: 257 reads here */omšo/, a reconstruction on which Janhunen 1993: 171–172 also agrees; as is well known, there is no cognate for this number in WM, *cf.* *juwan emu* '11', namely *juwan* '10' + *emu* 'one', though it is most likely related to the word *omšon* in the expression *omšon biya* 'the eleventh month'). Interestingly enough, Kane considers that †šomi contains the default Chinese character to transcribe the sequence /šo/. In consequence, he ignores the

¹ The Chinese transcriptions of the Late Jurchen items are quoted according to Kiyose [= H], Jīn (1984), and Kane [= A]. Early Mandarin reconstructions, between brackets, are cited according to Pulleyblank (1991). The actual readings of the Jurchen items (marked †) are presented in the traditional Romanization for Written Manchu.

ones used in *ᠳᠰoro* and *ᠳᠳoᠰo* (A 127). It is my understanding that the correct situation is just the opposite one: *ᠳᠰomi* must be dealt with as an exception, while words such as *ᠳᠰoro* and *ᠳᠳoᠰo* reflect the actual rule with regard to the transcription of the sequence /šo/. Since we have no more instances containing the character 說 to rely on, and H and Manchu unambiguously reflect /u/, I propose just to emend Kane’s reading for *ᠳᠰumi*.²

What could be the explanation behind such a particular orthographic rendition, I cannot tell. The lowering of /u/ to /o/ is a common feature in the Manchurian area, having been diachronically described in the domain of both Tungusic and Mongolic languages, even in Korean (see Janhunen 1981). As for Tungusic, see among others the systematic results in (Literary) Nanay /o/ & /u/ for the PT vowel pairs */o u/ & /ö ü/. However, this process is usually inserted in major evolutive systems (e.g. vowel rotations or shifts), therefore to invoke it to account for sporadic instances like this seems pretty frivolous. The lack of testimonies from Old Jurchen here is an infelicity, since this could help us to discern whether the initial sequence actually was */ö/ or */ü/ and in this way to explain why the Chinese transcription is somehow irregular. As is well known, Old Jurchen */ö/ and */ü/ yielded, depending on the context, either /u/ or /e/ in Late Jurchen (*vid. i.a.* Kiyose 1977: 41-43, 2000: 180-181).

Hypotheses in the direction of a kind of contamination with the Chinese word 深 (EM ᠰim) ‘deep’ would make little sense to me.

1.3.3. The segment */gi/ in the H form seems to correspond to WM +nggi (*vid. i.a.* Burykin 1999: 32). This suffix is unproductive in WM, therefore the words allegedly containing it must be dealt with as relics. Kiyose has recently proposed to read here */ŋi/ instead of */gi/ (2004: 98). Kiyose aims to reconstitute LJ */ŋ/ against the common assumption according to which PT & Old Jurchen */ŋ/ merged with */g/ already in LJ, at least in initial position (this was also Kiyose’s opinion until recently, *vid. i.a.* 2000: 179). Since the distinction is not vital for the purposes of the present paper, I shall not pursue the matter. Depending on the Jurchen character used, Jin (1984: 101) reads either */-gi/ or */-ŋgi/.

Alternating forms showing the presence or the absence of this element are very common even when LJ A is confronted with H, e.g. A 亦車 (ji tᠰ^hε) *ᠳice* ‘new’ [192] vs. H <儻受> 一車吉 (ji tᠰ^heki) *ᠳicegi* [666], cfr. WM *ice* id. Burykin mentions two additional instances in which LJ has */-gi/, while WM does not: A 塔兒麻吉 (t^ha^hᠮa^hki) *ᠳtalmagi* [16] ‘mist, frost’, H <儻幸僅> 塔馬吉 (t^ha^hma^hki) *ᠳtamagi* [018] vs. WM *talman* ‘mist, fog’ and A 塞忙吉 (sᠵma^hki) *ᠳsemani* ‘hoar-frost, rime’ [8], H <儻幸僅> 塞馬吉 (sᠵma^hki) *ᠳsaimagi* [009] vs. WM *sungke-* ‘to become hoary with frost’, see also Bikin Nanay (= Kilen) *sunḡu* ‘hoar-frost’ or (Eastern & Southern) Ewenki *sinjksä*

² Strictly speaking, what seems a reading mistake usually is nothing but the mechanistic interpretation of Chinese transcriptions. However, it is well known that Chinese transcriptions of foreign words are often impressionistic and/or distorted by later confusions, so one should not trust them altogether. This concerns not only Jurchen, but also other languages with the same material particularities, e.g. Khitan. Therefore, eventual emendations to Jurchen’s readings should be welcome, especially when they are based on (unproblematic) comparative data as is the case, rather than condemned or disliked.

id. (< Common Tungusic [= CT, i.e. with no Manchuric testimonies] */süŋü/, see TMS II.90b-91a). As for the latter, already Grube (1896: 98b, repeated by Kane) compared it to WM *silenggi* ‘dew’, but this word can be safely related to another set of words, all derived from PT */siili(.ksä)/ id. (see TMS II.85b-86a for cognates). It is worth noting that *†šumigi* and *†icegi* have the element */gi/ written with the Jurchen character <受>, while *†talmagi* and *†saimagi* show <僅>. Jīn reads *†šumigi*, *†icegi*, *†talmangi* and *†saimangi*, respectively. Kiyose (2004: 98) unfortunately does not address this problem. I believe, however, that this “scribal practice” may be somehow related to the fact that we are dealing here with two different suffixes that became homophonous only at a later stage. I will elaborate further on this issue somewhere else.

All in all, on the basis of instances such as those, the pair A *†šumi* vs. H *†šumi(ŋ)gi* turns out to be totally natural. The very same remark deserves the presence or absence of word-final /-n/ in the Manchu-Sibe and A forms, respectively. Although one could say that this phenomenon represents a Trans-Eurasian feature, i.e. the instability of the element /-n/ can be observed in Mongolic and Tungusic (see Vietze 1969 for further details), in this case the solution is much simpler: A systematically ignores this segment.

2.1. As far as the ultimate origin of Manchuric */šumin/ ‘deep’ is regarded, there are two major hypotheses.

2.1.1. According to the first one, which has two versions, WM *šumin* is related to Pan-Ewenki & Udihe & Oroch *suŋta* (but Nerchin Ewenki *suunkta*, EEWTD 750 [10114]), Xajlar Solon *süüŋta*, Literary Ewen *xuunta*, Arman *suunta*, Orok *sukta*, Literary Nanay *soŋta* (cfr. Maximowicz’s *suŋta* ~ *suŋkta* in Grube 1900: 96a), etc. (see TMS II.128a-b). Benzing, the author of this hypothesis (TSVG 37), proposed to reconstruct PT */suŋta/ ‘deep’. He seemingly assumed that the CT word can be segmented */suŋ-ta/. He does not specify what */-ta/ stands for. One can only speculate that he was referring to the adjectival suffix */+ta(a)/ (< */+tay/?), see TSVG 89 §104a), apparently unproductive already in PT. The resulting stem */suŋ-/ would have continued in Manchu as a derivative with the help of the PT suffix */±min/ (TSVG 90 §104f) as in *jiramin* ‘thick’, *golmin* ‘long’, etc., after which nasal reduction */-ŋm-/ > /m/ would have taken place. This proposal is not accepted in TMS (in consequence, the Manchu word is glossed apart, see TMS II.429b). Kiyose (2004: 98) accepts it uncritically. Though hesitant, Doerfer also accepts Benzing’s idea (EEWTD 750 [10116], “dies dazu gehörig?” about the Manchu word), but he proposes an even odder path of evolution to reconcile both the Manchuric and the CT data: (1) CT */sunta/ < */sumta/ < */syum-ta/ ~ */syum-cta/ > (2) CT */suŋ(k)ta/.

2.1.2. In its present formulation, Benzing’s proposal would have to be rejected on phonological grounds: (a) CT */s-/ does not correspond to PM */š-/ and consonant clusters of the type */-ŋN-/ are usually reduced to */ŋ/ with no trace of the second nasal element or just preserved untouched. Doerfer managed some of these problems. The glide component in Doerfer’s sequence */syu/ might account for the initial palatal /š-/ in Manchuric. However, the reconstruction of such a sequence opens a well known debate in Tungusic historical linguistics with regard the existence and ulterior evolution of the so-called diphthong(oid)s.

Benzing proposed that [ɳ̃] and [ʃ̃] were in the historical languages the outcome of PT */syV/ and */nyV/. At the same time, he acknowledged the possibility that such sequences could be reinterpreted as single (short or long) vowels preceded by palatalized consonants (see basic statements in TSVG 25, 40, 42 §§32, 52, 54). This is partially the stand taken by some other authors such as Doerfer or Cincius. They deemed unnecessary the reconstruction of diphthongoids, with the notable exception of */ya(a)/, supported the existence of PT */ñ/ (Cincius 1949: 190–194, 213–214 §§51–53, 65) and admitted that there is no PT */š̃/ but only */syV/. Thus, Literary Nanay *ñoŋña* ‘goose’ could be traced back, depending of the system one adopts (*pro* Benzing or *pro* Doerfer-Cincius), to PT */nyuŋnya(+kii)/ or */ñuŋña(+kii)/ id. (cfr. Pan-Ewenki *ñuŋñakii*, Orok *nuŋna*, Xajlar Solon *nunnaxi*, Literary Udihe *ñuŋñ'ai*, WM <niyongniyaha> = /ñoŋña.ha/, etc., see TMS I.611a, 646b–647a). Doerfer, who accepts the idea of diphthongoids under very specific conditions (no sequences */nyV/, but */ñV/, *vid. i.a.* 1978: 101–102),³ proposes to reconstruct */syu/. Be that as it may, to adopt one or another reconstruction is not an easy task, especially when the details concerning them have not been worked out yet. It is my understanding, however, that the only possible reconstruction with regard to /s/ is the one containing diphthongoids, otherwise we would have to assume that in some cases Literary Ewen or Northern Ewenki /h/ comes from */š̃/, an extremely odd and awkward sound change. I propose that diphthongoids could have merged with the corresponding long vowels as happens in some modern Tungusic languages (note that this decision would render technically impossible to guess whether the vowel of the diphthongoid was long or short), blocking the ulterior palatalization of the preceding consonant. Sporadic instances of palatalization are to be expected, however. Thus, basing on (Eastern & Southern) Ewenki *seen*, Xajlar Solon *šeeⁿ*, Negidal *seen*, Orok *see*, Literary Nanay *sean*, Literary Ewen *heen* ‘eye’, etc. (TMS II.70b–71b), I would reconstruct */sya(a)n/ *pace* Benzing or Doerfer’s */syaan/, i.e. my reconstruction cannot tell about vowel length, while Benzing and Doerfer only reconstructed long vowel (note Literary Ewen *h-* from *s-* and the sporadic palatalization in Xajlar Solon).

2.1.3. Instead of the complicated chain of facts proposed by Doerfer for WM, we may just assume that the adjectival suffix */-min/ was added to the resulting stem */syum-/, after what the simplification of the geminated consonant */mm/ is only natural, i.e. Proto-Manchuric */š̃um+(m)in/ > *šumin*. However, one important question remains unsolved: what is the exact role of the elements */-tA/ and */-ktA/ in Doerfer’s formulation. The exact identification, however, is a matter of greater speculation. The following suffixes are obvious contenders:

³ As Doerfer, I also support the reconstruction of PT */ñ/ instead of sequences containing diphthongoids. Typologically speaking, PT phonetic configuration would resemble very much the one of languages such as (Standard) Spanish: existence of diphthongs (and triphthongs) alongside the palatal subsystem /ñ ɰ ʃ̃/ and the fricative /s/. Then, I reconstruct diphthongoids (there are triphthongs only in some languages, e.g. Literary Nanay, and they are always secondary, *vid. i.a.* Janhunen 1985: 109–112) plus */ñ j c s/ for PT. Note that in Standard Spanish /s/ & /n/ are never automatically palatalized before /i/, i.e. /si(V)/ & /ni(V)/ are never realized [ši(V)] & [ñi(V)]. It is possible that the very same scenario should be set up for PT.

(1) The deverbal nominal suffix */-ktA/. It is most likely related to the same element in forms such as Pan-Ewenki *oomakta* ‘new’, from *oo-* ‘to become’ (the derivative suffix $\pm mA$ is used to create adjectives, e.g. *giramna* ‘bone’ → *giramna+ma* ‘(made of) bone’, *sumala-* ‘to be silent’ → *sumala-ma* ‘taciturn, silent’, etc., see Konstantinova 1964: 111). This suffix, already unproductive in most Common Tungusic languages, is used to derivate nouns expressing the result of the verbal action, e.g. Literary Ewenki *waa-* ‘to kill’ → *waakta* ‘wound’, *guun-* ‘to say’ → *guu(ni)ktä* ‘word’, etc. (Konstantinova 1964: 88). Konstantinova notes (*ibid.*, 113–114) that several derivative suffixes seem ambivalent with regard to the noun-adjective characterization of the output. Konstantinova includes the resultative */-ktA/ among them. This suffix is most likely related to the resultative participle suffix */-ktA/ (TSVG 119–120 §130f “iterativer Aspekt”). The derivative suffix was passed over in silence in Benzing’s comparative grammar because materials at his disposal did not contain a description of it, or one very vague, so it is only natural that Benzing would not devote to it much time. Actually, the connection is still largely ignored, and mentions to it occur only occasionally. Yet, many authors seem to recognize it. For example, in their description of Bikin Udihe, Nikolaeva and Tolskaya are aware of the non-productive suffix $\pm ktu$ creating both nouns and adjectives such as *koŋo-ktu* ‘lean, thin’, *sokto-ktu* ‘drunk’ or *ñä-ktu* ‘rotten’ (2001: 165) which they correctly, though hesitantly, relate to the resultative participle *-ktu* (*ibid.*, 197) via a “participle-to-adjective” conversion. While the details of this comparison await further research, I think the link between the nominal and the verbal */-ktA/ is safe enough to be endorsed.

(2) The collective suffixes */+tA/ and */+ktA/ (TSVG 71-72 §80). Adjectives in the Tungusic languages are divided into two categories on the basis of certain morphological characteristics: descriptive and relative, or non-relational and relational (or “proprietary”, a term borrowed from Australian linguistics, see Nikolaeva 2008: 970). Only the first type is of interest to us now. They are classified in two groups: derived and non-derived. In Literary Ewenki as well as in the rest of Tungusic languages, the latter class traditionally includes *aya* ‘good’, *xägdi* ‘large’, *gugda* ‘high’ or *suŋta* ‘deep’. Interestingly enough, these adjectives may be also nouns: *aya* ‘goodness’, *xägdi* ‘size’, *gugda* ‘height’ and *suŋta* ‘depth’ (*vid. i.a.* Lebedeva, Konstantinova, Monaxova 1979: 89–91, note Vasilevič talks about a “special group of nouns” in ERS 702–703). Could they, or some of them, be originally collectives which by some unknown reason semantically developed an adjectival side?

Irrespective of the fact that I have no answer for the question posed in (2), it is my understanding that option (1) offers a more convincing solution. As far as CT is regarded, */syum-/ was originally a verbal base meaning ‘to be deep, to deepen’ with a regular nominal derivative */syumkta/ ‘depth; deep’. It is worth noting that there seems to be no need to assume the existence of two different suffixes as Doerfer did. On the contrary, we only need to reconstruct one form, namely */syum-ktA/, the diversity of forms in the historical languages being the result of the way the /-Nkt-/ cluster was simplified, i.e. with (/ŋ(k)t-/) or without (/n(k)t-/) trace of the velar stop element. This actually is what the testimonies of Pan-Ewenki vs. Nerchin Ewenki and Literary Nanay

vs. Maximowicz’s Nanay seem to suggest. Note that vowel length could come from the merging of the diphthongoid with the long vowel /uu/ or as compensatory lengthening after the simplification of the consonant cluster */-ŋkt-/. Yet, vowel length remains unaccounted at least in Ewenki and, instead of Literary Nanay *soŋta*, one could ideally expect **seoŋta* <сионгта>. Unfortunately, I have no way to explain these two irregularities out.⁴

2.2.1. The second hypothesis, formulated in EDAL (II.1342-3), links WM *šumin* with (Eastern & Southern) Ewenki *comko* ~ *como* ‘scoop’ and Literary (= Xor) Udihe *compo-* ‘to submerge’ (TMS II.406a). The form *cumku-* ‘to pour water on oneself’ (TMS II.414a) is said in TMS (and consequently in EDAL) to be “Ewen”, but this actually is a clerical error. Abbreviations in TMS for Ewen and Ewenki are “Эвех.” and “Эвехк.”, respectively, so it is easy to confuse both languages if one makes a slip with the last “к”, as has happened in this case. The form *cumku-* must be assigned to the Northern Ewenki dialect of Ilimpi (cfr. ERS 529a s.v. *cumku-mii*). Be that as it may, these words share a common stem which at first sight would seem more suitable to account for the origin of the Manchu word under discussion.

2.2.2. There are, however, several reasons preventing us from accepting this proposal: (1) there is already a Manchu cognate for the set of Tungusic words quoted above, namely *côman* ~ *coman* ~ *como* ‘cup, goblet’, (2) the sound correspondence CT */c-/ vs. PM */š-/ is irregular, and (3) the semantic motivation behind those words, though related by the idea of deepness, may well belong to totally different semantic fields. All this being said, let’s take a look at each of these points:

(1) Curiously enough, the authors of EDAL mention neither these Manchu words nor Xajlar Solon *somo* ‘cup’, in spite of the fact that both WM and Solon data are included in the TMS *lemmata* which the authors of EDAL use as their main source. Additionally, all these Tungusic words are considered to be, on very solid bases, of Mongolian origin (cf. Written Mongolian *čomu* [čomo] ‘wine cup, goblet’, Khalkha *tsomoo*, id.vid. i.a. Doerfer 1985: 61 [141], MEM 50). The derivative suffix **/-kO/*, present in some Ewenki forms, is not described in any grammar, to the best of my knowledge. Instead, these forms may have arisen after contamination with native terms such as Pan-Ewenki *soko-* ‘to scoop, ladle’ (TMS II.105b). This explanation would also account for Literary Udihe *compo-*, if it is assumed that the original **/k/* assimilated to the previous bilabial consonant, i.e. **/mk/* > */mp/*. However, it is necessary to note that the Udihe cognate of Pan-Ewenki *soko-* has no *-k-* as expected, since it has been lost between vowels, see Literary Udihe *s’ou(n-)* ‘scoop, bucket’ < **/soko.wun/* or Samargin Udihe *s’olo-* ‘to scoop’ < **/soko.lo-*. It follows that the contamination would have taken place no later than the Pre-Udihe

⁴ At this point it is legitimate to ask whether it makes sense at all to include Manchu *šumin* here. If it would be excluded, there is no problem, one would argue, to reconstruct CT **/suŋta/*. Unfortunately, this option would still leave unanswered the presence of vowel length in Literary Ewen, Xajlar Solon or Nerchin Ewenki. We must bear in mind that PT vowel length is reconstructed on the basis of Literary Ewen, while Literary Nanay (generally speaking, all the Amur Tungusic languages) shows an extremely irregular pattern as far as the retention of this feature is regarded (just a cursory examination of Benzing’s instances in the section devoted to long vowels can well prove it).

stage. Incidentally, the Literary Nanay cognates of Pan-Ewenki *soko-* are *soopo*ⁿ ‘scoop’ & *soopo.la-* ‘to scoop’ (cfr. Kili *soko*ⁿ & *soo.lo-*, Kilen *soaŋko* & *sofo.la-* id.). These may be also the result of contamination, since they are clearly irregular if we compare them directly with Pan-Ewenki *soko*, etc. If we assume that */somko/ is the departure point for the Literary Nanay (and the Kili & Kilen) forms, then it is much simpler to explain vowel length (compensatory lengthening after the simplification of the */mk/ > */mp/ cluster in Literary Nanay, Ø < /-w-/ < /-p-/ in Kili?) and Kilen *soaŋko* < */somko/. All in all, we would have to posit two different words, namely PT */como/ and CT */soko/, with identical, or very similar, meanings (‘(to) scoop, bucket’) which in the course of time got mixed, yielding different, sometimes confusing, results in the historical languages.

(2) Norman discussed several decades ago (1977) that PT */t-/ yielded Manchu *s-* when in Proto-Manchuric */t-/ is followed by */-j-/ in the same stem, irrespective of the origin of the latter (i.e. PT */-j-/ , */-di(V)-/ or any of the consonant clusters that underwent such an evolution). For the intermediate stages of this process, the change */t-/ > */c-/ is described as the result of assimilation and */c-/ > (*/š-/ >?) /s-/ as deaffrication. The most appealing example is PT */tuñja/ ‘five’ > Proto-Manchuric */cunja/ > LJ A H < 𑮎𑮏 > 順劄 (šyn tšã) ʃšunja [1113, 640] (cfr. WM *sunja*, Sibe *sunjaa*), but there are more instances (see Norman 1977: 232–233). Generally speaking, the status of the process of deaffricating palatal affricates over the Eurasian territory is still disputed (for instance, the case of /c/ > /t/ has been described by some authors as a diachronically active sound change in some Uralic, Yeniseian, Tungusic, and Turkic languages, *vid. i.a.* Anderson 2003: 12–20). One could argue that the second step, i.e. */c-/ > /s-/ could actually be regular even when the departure point is PT */c-/. However, this is far from being the case because there are no instances reflecting the alleged sound change, i.e. */c/ > /š/, with Manchu *šumin* and Ewenki *como* ~ *comko* being the only example (*testis unus, testis nullus*).

However, as argued in the previous section, Ewenki and the rest of words related to this are of Mongolian origin. Doerfer claimed (1985: 177–179 and MEM 228, seemingly accepted by Janhunen in 1993: 172) that the sound correspondence /c/ vs. /š/ is a diagnostic feature reflecting an ancient layer of Mongolian loanwords in WM, e.g. Written Mongolian *nicügin* ~ *nicügen*, Buriat *nüsege*ⁿ &c. ‘naked’ → WM *nišehun* ~ *nišuhun*. On this basis, one could assume that both WM *côman* (*et alia*) and *šumin* correspond to different layers of Mongolian borrowings. This is a common phenomenon which may be easily corroborated with parallels from any given language around the world, e.g. Polish *skiba* ‘ridge’, *szyb(a)* ‘(window) pane’, *szajba* ‘washer’ reflect the different stages in the evolution of the very same item, i.e. Old High Germn *scība* > Middle High German *schībe* > German *Scheibe* ‘slice, disk, pane’, respectively (Majtczak 2010: 124). Though very attractive, the “multiple loanword” (or perhaps “updating loanword”?, see Majtczak 2010: 123 ft. 1) hypothesis still raises some questions. Sound correspondences involving vowels are an insurmountable obstacle, and semantics, though similar, is not convincing enough. Last but not least, I cannot elaborate further on what Mongolian language could have acted as the donor in the case of WM *šumin*.

(3) In case we are too lax about the precision of semantics, then many other words should be included in this etymological exercise. See for example Manegir and Nerchin Ewenki *čuman* ‘big birch basket for storing seeds’ (EEWTD 197 [2241]), said to be a Russian loanword. In its turn, the ultimate origin of Russian чуман (ESRJa IV.382) is still disputed (ESRZVS 693). There can be no question that these and some other words are linked to many of the forms quoted under the “Turkic” heading by the authors of EDAL (in the same *lemma* discussing the Tungusic materials mentioned above). As for their origin, already Clauson states in very clear terms that those Turkic words come from Persian *čamča* (EDT 422a-b, for the rationale see TMEN III.95–99 [1121]). Regrettably, the authors of EDAL do not elaborate on this point (it seems they ignore all borrowing allegations in order to present a case for an Altaic etymology). It is not my intention to attempt to deal with all these forms because firstly I lack expertise in some of the required fields, and secondly I do not think it would help to clear up the question regarding the origins of WM *šumin*.

3. Summing up, WM *šumin* ‘deep’ together with its derivatives and the rest of Manchuric cognates (LJ, Sibe) may share a common origin with other Common Tungusic words via PT */syum±/ ‘depth; to (be) deep(en)’. This idea was firstly suggested by Benzing, and later expanded by Doerfer. Building on Benzing’s and Doerfer’s original statements, I have proposed that Manchu *šumin* might be a derivative, i.e. */syum+min/, and that the historical continuations documented in the Common Tungusic languages should have departed from */syum-ktə/, with divergent results due to the simplification of the cluster /kt/ and its effects on the previous nasal. The Manchuric words seem to stem from a noun base, while in our analysis only a verbal base can be postulated for CT. Note, however, the ambiguity of the suffix */±min/ with regard to the nominal or verbal nature of the base to which it is attached. If not because of this minor detail, we could safely reconstruct a pure verbal base. The only negative aspect of this proposal is the distribution of vowel length, which, as in many other instances, is irregular.

Though there are numerous look-alike words (both Tungusic and non-Tungusic) which could seem to be related, I am very dubious about the validity of the potential links, regardless of whether this is genealogical or via borrowings.

References

- A = D. Kane, *The Sino-Jurchen Vocabulary of the Bureau of Interpreters*, Bloomington 1989.
 CDSM = K. Yamamoto, *Manshū-go kōgo kiso goi shū*, Tōkyō 1969.
 CMEL = J. Norman, *A Concise Manchu-English Lexicon*, Seattle-London 1978.
 EDAL = A.V. Dybo, O.A. Mudrak, S.A. Starostin, *An Etymological Dictionary of the Altaic Languages* [3 vols.], Leiden-Boston 2003.
 EDT = G. Clauson, *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*, Oxford 1972.
 EEWTD = G. Doerfer, *Etymologisch-ethnologisches Wörterbuch tungusischer Dialekte (vornehmlich der Mandschurei)*. Unter Mitwirkung von Michael Knüppel, Hildesheim-Zürich-New York 2004.
 ERS = G.M. Vasilevič, *Ėvenkijsko-russkij slovar’*, Moskva 1958.

- ESRJa = M. Fasmer, *Ètimologièeskij slovar' russkogo jazyka (perevod s nemeckogo i dopolnenija O.N. Trubačëv)*, t. 4, Moskva 1973.
- ESRZVS = A.E. Anikin, *Ètimologièeskij slovar' russkix zaimstvovanij v jazykakh Sibiri*, Novosibirsk 2003.
- H = G. Kiyose, *A Study of the Jurchen Language and Script*, Kyoto 1977.
- HM = E. Hauer, *Handwörterbuch der Mandschusprache*. Wiesbaden 1952–1955.
- MMC = H. Kuribayashi & Hurelbator (eds.), *Manchu-Mongolian-Chinese Triglote Dictionary of 1780 Arranged by Manchu Words*, Sendai 2008.
- MEM = W. Rozycki, *Mongol Elements in Manchu*, Bloomington 1994.
- MSM = J. Kim, D. Ko, D.O. Chaoke, Y. Han, L. Piao, B.V. Boldyrev, *Materials of Spoken Manchu*, Seoul 2008.
- PMRS = I. Zakharov, *Polnyj man'čžursko-russkij slovar'*, Sanktpeterburg 1875.
- TFAXB = C.H. Melles C.H. 1976. *Un vocabulaire du mandchou préclassique. Le Tongki fuqa aqû xergen-i bitxe*. "Acta Orientalia Academiae Hungaricae" 1976, 30/2, pp. 209–243.
- TMEN = G. Doerfer, *Türkische und mongolische Elemente im Neupersischen* [4 vols.], Wiesbaden 1963–1975.
- TMS = V.I. Cincius (ed.), *Sravnitel'nyj slovar' tunguso-man'čžurskikh jazykov* [2 vols.], Leningrad 1975–1977.
- TSVG = J. Benzing, *Die tungusischen Sprachen. Versuch einer vergleichenden Grammatik*, Wiesbaden 1956.
- Anderson G.D.S, *Yeniseic languages from a Siberian areal perspective*, in: E. Vajda, G.D.S. Anderson (eds.), *Studia Yeniseica: Typological Studies on Yeniseic (= Sprachtypologie und Universalienforschung 56/1-2)*, Berlin 2003, pp. 12–39.
- Burykin A.A., *Morphological aspects of the language of the Jurchen script*, in: J. Janhunen, V. Rybatzki (eds.), *Writing in the Altaic World*, Helsinki 1999, pp. 29–39.
- Cincius V.I., *Sravnitel'naja fonetika tunguso-man'čžurskikh jazykov*, Leningrad 1949.
- Doerfer D., *Urtungusisch ö*, in: W. Weiers (ed.), *Tungusica. Beiträge zur nordasiatischen Kulturgeschichte*. Wiesbaden 1978, pp. 66–125.
- Doerfer D., *Mongolo-Tungusica*, Wiesbaden 1985.
- Grube W., *Die Sprache und Schrift der Jučen*, Leipzig 1896.
- Grube W., *Goldisch-deutsches Wörterzeichniss mit vergleichender Berücksichtigung der übrigen tungusischen Dialekte* [= L. von Schrenck (ed): *Reisen und Forschungen im Amur-Lande*, vol. 3: *Linguistische Ergebnisse*]. St. Petersburg 1900.
- Janhunen J., *Korean vowel system in North Asian perspective*, "Han-geul" 1981, 172, pp. 129–146.
- Janhunen J., *On the Diphthongs in Nanay*, "Bulletin of the Institute for the Study of North Eurasian Cultures" 1985, 17, pp. 103–115.
- Janhunen J., *The teens in Jurchen and Manchu revisited.*, in: U.-M. Kulonen (ed.), *Festschrift für Raija Bartens*, Helsinki 1993, pp. 169–184.
- Jin Q., *Nǚzhēn wén cídiǎn*. Běijīng 1984.
- Kałużyński S., *Die Sprache des mandschurischen Stammes Sibe aus der Gegend von Kuldscha*, Warszawa 1977.
- Kiyose G.N., *Genealogical Relationship of Jurchen Dialects and Literary Manchu*, "Central Asiatic Journal" 2000, 44/2, pp. 177–189.
- Kiyose G.N., *Independent Corroboration of the Jurchen *ŋ- Reconstructed by the Comparative Method*, in: C. Naehrer (ed.), *Proceedings of the First International Conference on Manchu-Tungus Studies*, vol. 2, *Trends in Tungusic and Siberian Linguistics*, Wiesbaden 2004, pp. 93–99.
- Konstantinova O.A., *Èvenkijskij jazyk. Fonetika i morfoložija*, Moskva–Leningrad 1964.
- Lebedeva E.P., O.A. Konstantinova, I.V. Monakhova, *Èvenkijskij jazyk. Posobie dlja pedagogičeskikh učilišč*. Leningrad 1979.

- Majczak T. *Gra japońskimi kartami (przyczynek do teorii badań kontaktów językowych)*, „Biuletyn Polskiego Towarzystwa Językoznawczego” 2010, 66, pp. 123–132.
- Nikolaeva I., M. Tolskaya, *A Grammar of Udihe*, Berlin 2001.
- Nikolaeva I., *Between nouns and adjectives: A constructional view*, “Lingua” 2008, 118, pp. 969–996.
- Norman J., *A Sketch of Sibe Morphology*, “Central Asiatic Journal” 1974, 18, pp. 159–174.
- Norman J., *The evolution of Proto-Tungusic *t to Manchu s*, “Central Asiatic Journal” 1977, 21, pp. 229–233.
- Pulleyblank E. *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver 1991.
- Vietze H.P. Plural, *Dual und Nominalklassen in altaischen Sprachen*, “Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin” 1969, 18/3, pp. 481–512.