



Universiteit
Leiden
The Netherlands

A grammar of the Thangmi language with an ethnolinguistic introduction to the speakers and their culture

Turin, M.

Citation

Turin, M. (2006, May 17). *A grammar of the Thangmi language with an ethnolinguistic introduction to the speakers and their culture*. Retrieved from <https://hdl.handle.net/1887/4458>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4458>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/4458> holds various files of this Leiden University dissertation.

Author: Turin, Mark

Title: A grammar of the Thangmi language with an ethnolinguistic introduction to the speakers and their culture

Issue date: 2006-05-17

PART 1

GRAMMAR

CHAPTER 1

THE LINGUISTIC CLASSIFICATION OF THANGMI

1. Early classifications of Thangmi within Tibeto-Burman

The three-page grammatical description of Thangmi, then referred to as ‘Thāmi’, in the *Linguistic Survey of India* compiled by George Abraham Grierson, does not begin auspiciously:

The Thāmis have formerly been considered to speak the same dialect as the Sunwārs. During the preparatory operations of this Survey the two dialects were confounded in Darjeeling, and separate returns were only made from Sikkim. (1909: 280)

Sten Konow, the author of this passage, concludes his introduction on a more promising note when he states that Thangmi is actually ‘quite distinct from Sunwār’, and that despite being ‘much influenced by Aryan dialects’, it appears to be ‘a dialect of the same kind as Dhīmāl, Yākhā, Limbu, etc.’ (1909: 280). This description appears in Volume III, Part I of Grierson’s *Survey*, in a section entitled ‘Eastern Pronominalized Languages’. Thangmi was then classified alongside Barām (then referred to as ‘Bhrāmu’) as forming an ‘Eastern Subgroup’ of the ‘Complex Pronominalizing’ branch of ‘Himalayan Languages’ within the ‘Tibeto-Burman’ language family (1927, Vol. I, Part I: 58). Konow based his putative classification on word lists collected by Brian Houghton Hodgson half a century earlier, specimens of which he provided in the publication. Hodgson himself had recorded these languages as ‘Thāmi’ and ‘Bhrāmú’ respectively, although in the present context, ‘Thangmi’ and ‘Barām’ are more ethnolinguistically appropriate terms.¹

The words and phrases presented in Konow’s list were collected from Thangmi speakers in Darjeeling and make for interesting reading. The lexical items are considerably influenced by the Nepali language, as one might expect from linguistic data collected in the tea estates of north-east India where indigenous tongues were often jettisoned in favour of Nepali, the *Verkehrssprache* or vehicular

¹ The complex issue of the most suitable choice of ethnonym for the Thangmi people and their language is dealt with in depth in Section §2 of Chapter 2. According to George van Driem, the Barām call their language *Bāl Kurā*, the ‘language of the people’, in which *kurā* is Nepali for ‘language’ and *bāl* is Barām for ‘people, person, somebody’. While the term ‘Barāmu’ is allegedly still known to a few elderly non-Barām Nepali-speakers, the Barām themselves universally reject both ‘Barāmu’ and ‘Bhrāmú’, and insist on the use of the term ‘Barām’ in Nepali to describe the group and their language (van Driem 2001: 766).

‘language of commerce’. It is revealing that Thangmi words and phrases recorded in Darjeeling almost 150 years ago show a greater degree of Nepalification than contemporary Thangmi spoken in the districts of Dolakhā and Sindhupālcok in Nepal.²

In his *Introduction to Sino-Tibetan*, Robert Shafer adds his support to the Grierson-Konow proposition of a close genetic relationship between Thangmi and Barām by placing them together in the ‘Eastern Branch’ of the ‘West Himalayish Section’ of the ‘Bodic Division’ of ‘Sino-Tibetan’ (1974: 145). Following Shafer’s classification, Thangmi and Barām would therefore also be close relatives of other West Himalayish languages such as Byangsi, Manchad and Zhangzhung. Shafer admits that this classification is ‘tentative’, but is in no doubt that ‘Thami and Bhramu are closely related’ (1974: 145). Regarding their affinity to other West Himalayish languages, Shafer is similarly cautious: ‘From the limited vocabularies of them one can only say that they are here placed in West Himalayish because they appear to be closer to that group than [sic] to any other’ (1974: 3). While the empirical basis for Shafer’s hypotheses was scanty, his belief in a close linguistic relationship between Thangmi and Barām has been of more lasting interest than his classification of these two languages as West Himalayish.

Shafer posited nine lexical similarities shared by Thangmi and Barām which he believed indicated a degree of close genetic relationship (1966: 128). These nine lexical items are given in Table 1 below.

TABLE 1. SHAFER’S PROPOSED LEXICAL SIMILARITIES

Thami	Bhrámú	English
<i>dī-ware</i>	<i>dé</i>	one
<i>nis</i>	<i>ni</i>	two
<i>u-ni</i>	<i>u-ní</i>	sun
<i>tśala</i>	<i>chala-wani</i>	moon
<i>nem</i>	<i>nam</i>	house
<i>su-wa</i>	<i>s-wá</i>	tooth
<i>tśiya</i>	<i>chá</i>	eat
<i>ku-lna</i>	<i>ká-pá</i>	ear
<i>ka-pu</i>	<i>ká-pá</i>	head

² Examples are given in Chapter 2, Section §4.3.

Of the nine lexical correspondences, seven may now be discounted since they are either widely attested in other languages or easily reconstructed to Proto-Tibeto-Burman forms, leaving only two possible words supporting a special link between Thangmi and Barām. The comparative evidence is as follows: the Barām and Thangmi words for ‘one’ seem to derive from the Proto-Tibeto-Burman root **t(y)ik* ‘one’ (Benedict 1972: 94) or **tyak ~ *g-t(y)ik* ‘one, only’ (Matisoff 2003: 616), while the words for ‘two’ in both languages are also reflexes of the widely-attested Proto-Tibeto-Burman root **g-ni-s* (1972: 16) or **ʔ-nit ~ *ni* and **g/s-ni-s* ‘two’ (Matisoff 2003: 604). Consequently, the words ‘one’ and ‘two’ only indicate the already indisputably Tibeto-Burman nature of Barām and Thangmi, and do not indicate any special relationship between the two languages. Likewise, where Shafer suggests that Barām *s-wá* ‘tooth’ and Thangmi *su-wa* ‘tooth’ are unusual forms, both can now be reconstructed to the Proto-Tibeto-Burman root **s-wa* ‘tooth’ (Benedict 1972: 106) or **swa* ‘tooth’ (Matisoff 2003: 604), and Barām *chá* ‘eat’ and Thangmi *tšiya* ‘eat’ are similarly reflexes of the common Proto-Tibeto-Burman root **dza* ‘eat’ (Benedict 1972: 28) and **dzya* ‘eat’ (Matisoff 2003: 648). When Shafer suggests that Barām *ká-pá* ‘head, ear’ and Thangmi *ka-pu* ‘head’ are unique, he may have been unaware of the Nepali form *kapāl* ‘head, hair’ and the Kuvār form *ká-pá* ‘head’. Even in the little known language of Thochú, the form *kapat* ‘head’ has been attested (Hodgson 1880: pull-out section containing the *Comparative Vocabulary of the languages of Hôr Sôkyeul and Sifán*). It seems more plausible to suggest that the words for ‘head’ in both languages are Indo-Aryan loans rather than arguing for a separate lexicogenesis. Finally, the Thulung word *nem* ‘house, dwelling place’ (Allen 1975: 224) is cognate with Thangmi *nem* and Barām *nam*, both meaning ‘house’. All that remain are two lexical correspondences, Barām *u-ní* and Thangmi *u-ni* meaning ‘sun’ (perhaps both derived from **nəy* ‘sun, day’ as noted by Matisoff (2003: 604)), and Barām *chala-wani* and Thangmi *tšala* meaning ‘moon’ (both likely cognate with **s/g-la* ‘moon, month’ as reconstructed by Matisoff (2003: 599)). Most of the above data were carefully summarised by the Leiden linguist Arno Loeffen (1995), who reached the conclusion that Shafer’s evidence for grouping Thangmi and Barām together was at best based on two lexical isoglosses showing a specific phonological innovation.

Despite the paucity of empirical data for his classification, it appears from more recent research that Shafer’s suspicion of a special relationship between the two languages may indeed have been correct. The two proposed lexical isoglosses shared by Thangmi and Barām are now further supported by numerous morphological correspondences, particularly in the realm of verbal agreement affixes (van Driem, forthcoming). While the Barām system of verbal agreement has all but decayed, the

verbal morphology of Thangmi is complex and reminiscent of the Kiranti model. The completeness of the Thangmi verbal paradigm may even provide an insight into the degenerated Barām agreement system.

Six years after the publication of Shafer's *Introduction to Sino-Tibetan*, Paul King Benedict's *Sino-Tibetan: A Conspectus* was published. In this classic work, Thangmi and Barām are passed over without specific mention and are classified as belonging to what Benedict labels a 'Himalayish' grouping within 'Tibetan-Kanauri' (1972: 7).³ More important to the present discussion than the virtual absence of Thangmi and Barām, however, is Benedict's suggestion that although the Newar language could not be 'directly grouped with Bahing and Vayu [now Hayu]' (1972: 5-6), it nevertheless showed 'interesting lexical agreements' with them, and 'might be regarded as a Bodish-Bahing link' (1972: 8). The ambiguous position of Newar within Tibeto-Burman had also been noted by Shafer, who rejected Konow's typological classification of the language as 'non-pronominalised', but remained unsure of its genetic position.⁴ Shafer and Benedict's tentative exploration of a Newar-Kiranti link would lie dormant for some twenty years before being re-explored in George van Driem's *Mahākirāntī* hypothesis.⁵

2. Thangmi in light of the Proto-Kiranti verb

Kiranti languages are typically characterised by verbal agreement systems which are complex even by Tibeto-Burman standards. Conjugations of Kiranti verbs may have two or three prefixal slots and up to eight suffixal slots, and person-number agreement is frequently encoded through *portemanteau* morphemes or even tensed *portemanteau* morphemes, especially when involving a first person singular actant (van Driem 1990). It is generally accepted that the identification of slots or functional positions facilitates the comparison of cognate verbal morphologies, as the order of affixal morphemes in Tibeto-Burman verbal conjugations reflects the non-random sequencing of an ancient element order in the proto-language.

In Thangmi, an inflected simplex form consists of a verb stem to which affixes are attached, indicating tense and showing person and number agreement with

³ In his *Handbook of Proto-Tibeto-Burman*, Matisoff puzzlingly suggests that 'Thami' is part of the Chin subgroup of Tibeto-Burman (2003: 702).

⁴ Shafer wrote: 'From the limited number of comparisons brought together here one may tentatively say that Newarish (Newari and Pahri) is probably neither Baric nor Karenic, but somewhat intermediate between Bodic and Burmic; that is, its ties are with languages to the north (Tibet) and the east (Burma and the Indo-Burmese frontier) rather than with Tibeto-Burman languages of Assam' (1952: 93).

⁵ In the intervening years, Scott DeLancey described an 'Eastern Himalayan' grouping, which would include 'the Kiranti languages and others in eastern Nepal; probably also Newari' (1989: 321).

one or both of the actants of the verb. Unlike many Kiranti languages, however, Thangmi does not differentiate for dual number, nor does it exhibit an inclusive-exclusive distinction. Since the detailed workings of the Thangmi verb are analysed in Chapters 6 and 7, it will suffice for the present to discuss the verbal agreement affixes of the Thangmi conjugational paradigm in the context of what is known about the Proto-Kiranti verbal agreement system.

Previous comparisons of Kiranti verbal agreement systems show the conjugations of Kiranti verbs to reflect a split ergative pattern in which third person actants are marked differently than first and second person actants (van Driem, 1991b: 346). In Kiranti languages, markers indicating the involvement of a third person actant usually reflect the so-called ‘accusative system’ by which a third person patient (3P) and a third person agent or subject (3AS) are marked by a separate set of morphemes. On the other hand, markers denoting the involvement of a first or second person actant follow an ergative pattern: one set of morphemes indicates first or second person agent (12A) while another set denotes first or second person patient or subject (12PS). Moreover, number of actant is ‘indexed in the verb by different but apparently cognate morphemes for third person versus first and second person actants’ (van Driem 1991b: 346). As the synchronic morphemic analysis of the Thangmi verbal agreement system given in Chapter 6 demonstrates, Thangmi conforms to the Kiranti split ergativity model in structure while differing in the specifics.

Morphemes in the Thangmi affixal string offer an insight into the particular pattern of split ergativity in the language: the suffix <-i> denotes first person plural patient or subject (1pPS), reflecting an ergative agreement pattern. On the other hand, the zero morpheme (SAS) marking singular number of a second or third person (i.e. non-first person) agent or subject, the zero morpheme (3AS) marking the involvement of a third person agent or subject and the number suffix <-eŋ> marking plural number of a second or third person agent or subject (pAS), illustrate the accusative pattern. This accusative pattern is also reflected in Thangmi by the presence of the third person patient morpheme <-u> (3P). As this analysis demonstrates, Thangmi exhibits a pattern of split ergativity, similar but essentially different to that of the Kiranti type. It is also worth noting that the Thangmi verbal conjugation shows a significant level of morphological fusion reflected by a disproportionately large number of *portemanteau* suffixes, i.e. five out of seventeen.

Of the seventeen affixes in the Thangmi verbal agreement system, at least eight reflect the same morphological etyma as the corresponding Proto-Kiranti morphemes and are also cognate with morphemes attested in extant Kiranti languages spoken in Nepal. The Thangmi negative morpheme <ma-> (NEG) is

cognate with the Limbu and Dumi negative prefixes <me-> and <mə-> respectively, though negative prefixes in Kiranti are often tensed morphemes. The Thangmi reflexive suffix <-si> (REF) is cognate with Proto-Kiranti *<-nši> (REF), as well as with the Limbu and Bahing reflexive markers <-siŋ> (REF) and <-si> (REF) respectively. The Thangmi second person singular morpheme <-na> (2s) is cognate with Thulung and Lohorong <-na> (2s), both indexing the involvement of a second person singular, all of which are reflexes of Proto-Kiranti *<-na> (2s). Likewise, the Proto-Kiranti third person patient morpheme *<-u> (3P), has reflexes in many extant Kiranti languages such as Lohorong and Limbu <-u> (3P), as well as in Thangmi <-u> (3P). The /w/ in the Thangmi *portemanteau* suffix <-wa> (1p→23), marking the transitive relationship between a first person plural agent and a second or third person patient, may also derive from the Proto-Kiranti third person patient morpheme *<-u> (3P). The Thangmi second person plural morpheme <-ni> (2p) is a reflex of Proto-Kiranti *<-ni> (2p), and is further found in modern Kiranti languages such as Thulung, Lohorong, Kulung and Bahing in which <-ni> (2p) also indicates the involvement of a second person plural actant in a verbal scenario.

In Kiranti languages, as well as in the Tibeto-Burman family in general, the presence of a velar nasal /ŋ/ often indicates the involvement of a first person singular actant. As van Driem suggests, ‘most first-singular morphemes in modern Kiranti languages consist of the velar nasal /ŋ/ with some associated vowel preceding or following the nasal’ (1991b: 350). The Thangmi first person singular actant morpheme <-ŋa> (1s) is seen to be a reflex of either, or both, Proto-Kiranti *<-ŋ>, the first person singular agent marker (1SA), and *<-ŋa>, the marker of a first person singular actant in non-preterite time (1s/NPT). In extant Kiranti languages, reflexes of these Proto-Kiranti forms are found in Limbu <-ŋ> (1SA), Lohorong <-ŋa> (1s) and Dumi <-ŋ> (1s).

The Thangmi tensed *portemanteau* morpheme <-uŋ> (1s→3/PT), marking the transitive relationship between a first person singular agent and a third person patient in preterite time, is typically Kiranti in both form and function. The Thangmi tensed *portemanteau* morpheme <-uŋ> may be a fusion of the Proto-Kiranti morpheme *<-u>, denoting third person patient (3P), with the proto-morpheme *<-aŋ>, denoting first person singular actant in preterite time (1s/PT). If this analysis is accepted, with regard to the *portemanteau* morpheme <-uŋ> (1s→3/PT) at least, it appears that Thangmi is an extant example of the Proto-Kiranti model and arguably more canonically Kiranti in morphological structure than many of the other extant Kiranti languages. The Thangmi morpheme <-uŋ> (1s→3/PT) may also be cognate with Hayu <-ŋ ~ -N ~ -sŋŋ> (1s→3) or Hayu <-suŋ> (1sPS/PT), Limbu <-paŋ> (1s→3/PT) and Bahing <-ŋŋ> (1s→3/PT).

The Thangmi morpheme <-i>, denoting first person plural patient or subject (1pPS), is likely a reflex of the Proto-Kiranti inclusive suffix *<-i>. Thangmi makes no inclusive-exclusive distinction, and it is thus to be expected that the reflex of the Proto-Kiranti inclusive marker indicates the involvement of a plural first person in a verbal scenario, since both first person and plural number are implicit in any inclusive category. Moreover, when the Thangmi reflex is shown alongside reflexes of this proto-morpheme in other Kiranti languages, it becomes clear that the above extrapolation is in accordance with the data. In Lohorung, for example, the reflex of the Proto-Kiranti inclusive suffix *<-i> also denotes first person plural patient or subject (1pPS) and takes the form <-i>, while in Thulung the related form <-i> is attested, indicating the relationship between a first person plural inclusive agent and a third person patient (1pi→3).

The remaining morphemes present in the Thangmi verbal agreement system appear to be unrelated to their Proto-Kiranti counterparts, and while the affixes may index similar meanings, their surface forms are very different. Likewise, five Proto-Kiranti morphemes which have well-attested reflexes in extant Kiranti languages have no reflexes in Thangmi.⁶ It should be added, however, that in the case of three of these five Kiranti proto-morphemes, it is not that the categories which they encode are marked by other, non-cognate, morphemes in the Thangmi affixal string, but rather that they are not marked at all. Since Thangmi has no specific dual category, it follows that Thangmi will have no morphemes to mark dual agent or subject (dAS) or third person dual patient (3dP). Similarly, there is no specific third person plural agent (3pA) category in Thangmi, as the Thangmi morpheme <-eŋ> (pAS) marks only a second or third person plural agent or a third person plural subject.

As demonstrated above, it appears that Thangmi occupies a half-way house between a canonical Kiranti-style complex verbal agreement system and that of the less inflecting Tibeto-Burman languages. This conclusion was also reached some thirty years earlier by the French linguist Geneviève Stein, who correctly noted that the Thangmi speak a ‘pronominalized Tibeto-Burman language’ but hesitated to put it together with the Kiranti languages, because ‘although pronominalized, it does not present as complex a verbal morphology as these languages do, [nor] a proper dual nor an opposition inclusive/exclusive’ (as cited in Miller 1997: 116).⁷ While the link between Thangmi and the Kiranti languages will be apparent from the above

⁶ For a full analysis of the evidence, as well as charts of the Proto-Kiranti verbal agreement system as reconstructed by George van Driem, see Turin (1998a).

⁷ Stein never published her findings, and the citation provided by Miller derives from personal communication. Stein’s research findings are dealt with in greater detail in Chapter 2.

comparisons, the genetic position of Thangmi in relation to other Tibeto-Burman languages is discussed in the remaining sections of this chapter.

3. Before and after Mahākirāntī

At the 13th annual meeting of the Linguistic Society of Nepal, George van Driem advanced his *Mahākirāntī* or ‘greater Kiranti’ theory: a ‘hypothetical genetic unit’ including Kiranti and Newar (1992: 246).⁸ While his idea attracted both immediate attention and criticism, van Driem continued to refine his thinking as new linguistic data (specifically on Thangmi and Barām) came to light. In 2001, van Driem redefined the Mahākirāntī group as consisting of ‘the Kiranti languages proper and...the Newaric languages Newar, Barām and Thangmi. The set of languages which are related to Mahakiranti...includes Lepcha, Lhokpu and the Magaric languages’ (2001: 591).⁹ In *Languages of the Himalayas*, van Driem sets out the implications of his theory:

the linguistic ancestors of modern Mahakiranti groups and of Bodic language communities, which appear to be closer to Mahakiranti than to Bodish, peopled the Himalayas from the east and form a cluster of languages connected not only by shared geographical provenance but perhaps also related by more intimate genetic association and shared prehistorical contact situations. (2001: 590-591)

But what evidence did van Driem provide for the existence of the Mahākirāntī grouping? Dismissing lexical data as merely ‘suggestive’ and inadequate for ‘systematic comparison to yield decisive evidence’ (2003: 23), van Driem has stressed that the comparison of inflexional morphology provides evidence of a ‘highly sound and compelling kind’ (1992: 246). The morphological evidence of the Kiranti-Newar genetic link comes from Dolakhā Newar, the ‘most divergent...dialect of the language’ (van Driem 2001: 759) spoken in and around Dolakhā, an ancient Newar settlement and trading post ‘dating back perhaps as far as the Licchavī period [circa 300-879 A.D.]’ (2001: 759). The verbal agreement system of Dolakhā Newar is cognate with the conjugational morphology attested in Kiranti languages: verbs in the Dolakhā dialect of Newar agree for person and number with the intransitive subject and transitive agent in all tenses. Not only is the structure of Dolakhā Newar verbal

⁸ The term of choice in English for both the indigenous people and language of the Kathmandu valley is ‘Newar’, and emphatically not the Aryan-inspired ‘Newari’, which is generally considered to be offensive to contemporary Newar sensibilities.

⁹ That Magaric languages may be genetic relatives of Mahākirāntī is an interesting proposition. A different interpretation is offered by the Newar linguist, Tej Ratna Kansakar, who places Thangmi in a so-called ‘Magar Group’ of languages distinct from ‘Kirantish Languages’ (1993: 167).

morphology reminiscent of the Kiranti model, but Dolakhā also appears to be one of the more archaic and conservative of the extant Newar dialects. Van Driem makes this point succinctly:

Classical Kathmandu Newar...retains vestiges of a verbal agreement system like that of Dolakhā Newar. Therefore, the Classical Newar system is likely to derive from a more complete verbal agreement system, and the Dolakhā Newar verb probably represents a more faithful reflexion of this Proto-Newar system. (2001: 764)

More specifically, however, the Dolakhā dialect of Newar shares an important morphological trait with Thangmi and the Kiranti languages. In Dolakhā Newar, the morpheme <-u>, indexing third person future (3/FUT), is a verbal agreement suffix and also a reflex of the Tibeto-Burman proto-morpheme *<-u>. More specifically, the <-u> suffix in Dolakhā Newar denotes the involvement of a third person actant in the syntactic role of patient, a meaning also found in Thangmi and the Kiranti languages proper, as described in Section §2 above. As van Driem writes elsewhere:

The third person proto-morpheme *<-u> is ubiquitously reflected in Tibeto-Burman...In the Himalayas, these reflexes are all suffixes, and, in Kiranti languages, they all denote third person *patient* involvement. The Dolakha data likewise reflect third person patient marking: The vestigial suffix <-u> in the negative indicative, singular imperative and singular optative of r-stem verbs is clearly associated with grammatical patient marking, as it occurs only after transitive verbs. Similarly, in the past indicative, third singular subject is indexed by the suffix <-a> in intransitive verbs, but by <-u> in transitive verbs. (1993b: 36-37)

While acknowledging that verbal morphology constitutes only ‘one type of evidence which has yet to be corroborated by regular lexical and phonological correspondences’ (1992: 246), van Driem points out that the morphological evidence for the antiquity of the Dolakhā system is ‘decisive because in comparative linguistics conjugational agreement endings such as Dolakhā Newar *<-u> or the third person singular ending <-s> in the English present tense are precisely the type of elements...which are inherited, not borrowed’ (2001: 764-765).¹⁰

While reactions to the Mahākīrāntī hypothesis have been mixed, the strongest reaction against the proposed grouping came not from Western linguists, but from academics and lay people within the Newar community. Van Driem describes their resistance as ‘inherently suspect’ (2001: 599), pointing out that their unwillingness to accept the Mahākīrāntī hypothesis stems from the social exclusivity

¹⁰ It is prudent to note that flexional morphology is the heart of the inherited portion of any language, and genetic relationships between Indo-European languages had been firmly established on morphological grounds long before sound laws were discovered.

of the Newar community. Van Driem is, however, careful to acknowledge that while the communities may be linguistically related, in a ‘cultural sense these language communities could not be more different’ (2001: 599), and he points to the gulf in the socio-cultural worlds between the different Mahākīrāntī groups:

The Newars have for centuries had an advanced metropolitan culture, and, though they are linguistically Tibeto-Burmans, the Newars cultivated their own flourishing Sanskrit literary tradition. By contrast, the Kiranti, i.e. Rais and Limbus, were rural agriculturalists of the eastern hills, whereas the Barām and the Thangmi have remained amongst the socio-economically most disadvantaged groups of central Nepal. (2001: 599)

After a linguistic field trip to Bhutan in 2001, however, van Driem began to reconsider his Mahākīrāntī hypothesis. While in Bhutan, he collected data on the Gongduk language, particularly on its conjugational morphology and biactantial agreement system which contains reflexes of the Proto-Tibeto-Burman third person patient morpheme *-<u> (3/P).¹¹ On analysing the data, van Driem realised that:

the two specific morphological traits shared between Newar and Kiranti are not unique to Newar and Kiranti, but would appear to be the shared retention of a far older trait of the Proto-Tibeto-Burman verbal agreement system. Nothing else about Gongduk suggests any immediate affinity with either Newar or Kiranti within Tibeto-Burman. Therefore, the narrow but morphologically highly specific empirical basis for entertaining the Mahakiranti hypothesis no longer exists. (2003: 23-24)

In his conclusion to this article, van Driem suggests that while he no longer entertains the Mahākīrāntī hypothesis, the ‘case for Newaric or Mahānevārī has grown’ (2003: 25), and proposes that Thangmi and Barām ‘together form a coherent subgroup within the Tibeto-Burman family’ (2003: 24). Accordingly, the linguistic relationship between the Newaric languages (Newar, Thangmi and Barām) antedates ‘by a large margin the rise of the great Newar urban civilisation in the Kathmandu Valley, let alone the much later emergence in the XVIIIth century of the political entity of the kingdom of Nepal’ (van Driem 2001: 599).¹² In Section §4 below, I present a number of specific lexical isoglosses which further support the antiquity of the proposed Newar-Thangmi link.

¹¹ Van Driem draws attention to the Gongduk *portemanteau* suffix <-uŋi ~ -oŋe> (1→3) when compared with the first person subject morphemes <-ŋtji> and <-ŋni>, and to the Gongduk *portemanteau* suffix <-uri ~ -ore> (2p→3) when compared with the second person plural subject morpheme <-ire> (2003: 23).

¹² For a list of major Newar settlements which are believed to date back to the Kiranti period, see van Driem (2001: 732).

4. Thangmi-Newar lexical correspondences and the case for Newaric

4.1 Shared numeral classifiers

Following the clues suggesting a special relationship between Thangmi and Newar outlined in the first incarnation of the Mahākīrāntī hypothesis, I pursued the evidence for the proposed genetic link further. Supporting data came from the unlikely corner of a common set of numeral classifiers shared by the Sindhupālcok dialect of Thangmi and the Dolakhā dialect of Newar. A brief word about numeral classifiers in Tibeto-Burman languages will serve as a suitable point of departure.

Aside from the well-attested case of Newar, few of Nepal's Tibeto-Burman languages show any sign of having an involved numeral classifier system. A number of Kiranti languages do show remnants of classificatory systems, however, the best known instance of which comes from the pioneering study of Thulung by the Oxford-based anthropologist Nicholas Allen. Allen reports that in 19th century Thulung, as studied by Hodgson, 'countable nouns fell into classes defined by classifier particles associated with numerals' (1975: 113). Allen isolated six classifying particles (CLF) for Thulung: <-bop> meaning 'round objects' (or 'rounds' in Hodgson's notes), as in *ko bop miksi* (one CLF eye) 'one eye'; <-seol> meaning 'elongated object' as in *ko seol khel* (one CLF leg) 'one leg'; <-phe> meaning 'flat object' as in *ko phe nophla* (one CLF ear) 'one ear'; <-waŋ> meaning 'hollow circular object'; <-phu> meaning 'growing things, trees' and <-si> meaning 'holes, roads'. Allen goes on to describe what he calls significant 'variability' in the choice of particle, adding that this might indicate that 'the classifier system was beginning to break down' even in Hodgson's time (1975: 113-115).

More recent evidence of numeral classifiers present in extant Kiranti languages has been collected by members of the Himalayan Languages Project of Leiden University, corroborating Hodgson's early findings. For example, Joyce van Hoorn documents *sumbo?* 'three' in Chiling (personal communication), a fusion of the Tibeto-Burman numeral *sum* 'three' and a numeral classifier *bo?*, most likely cognate with Thulung <-bop> meaning 'round objects'. Similarly in Sāmpāñ, another Kiranti language, *i-bo* 'one' is made up of the numeral *i* 'one' and the classifier <-bo>, once again cognate with Thulung <-bop> meaning 'round objects' (René Huysmans, personal communication). Dumi also attests a numeral classifier cognate with Thulung <-bop>, in *mu-bo* 'six', segmented by van Driem as *mu* 'six' and the classifier <-bo> (1993a: 87-89). While interesting for comparative and historical reasons, however, these Kiranti classifiers have little in common with those attested for Thangmi or Newar.

Newar numeral classifiers, on the other hand, have received considerable attention from linguists of Tibeto-Burman languages and beyond, and a full discussion of the scholarship on this feature of Newar grammar is beyond the scope of the present discussion.¹³ In the following paragraphs, attention is focussed on the set of classifiers apparently cognate between Newar and Thangmi.

In her descriptive and historical account of the Dolakhā dialect of Newar, Carol Genetti notes that numerals are ‘always followed directly by numeral classifiers’ and describes ten classifiers which ‘are not used with any other nominal modification besides numerals’ (1994: 68). Seven of these classifiers are cognate with Thangmi numeral classifiers or nouns used in the Sindhupālcok dialect of Thangmi. In each case, the Newar and Thangmi classifiers are similar in both form and function.

The Thangmi noun *daŋ* ‘year’ from the Sindhupālcok dialect is likely cognate with the Dolakhā Newar classifier <-da> ‘years’ (Genetti 1994: 69), and the Thangmi classifier <-paṭe> ‘clothes, bamboo mats’ is probably cognate with the Dolakhā Newar classifier <-pta> ‘clothes (vests, pants, rugs, shirts, raincoats)’ (Genetti, personal communication). The Thangmi classifier <-pur> ‘branches, trees, long things’ may well be cognate with the Dolakhā Newar classifier <-pu> ‘hairs, bananas, ropes, necklaces, garlands, tongues, branches, sticks, brooms, pens’ (Genetti 1994: 69), and the Thangmi classifier <-pa> ‘leaves, paper, thin or flat things’ may be cognate with either the Dolakhā Newar classifier <-pat> ‘leaves, pieces of paper, silver leaf’ (Genetti 1994: 69) or the classifier <-pā> ‘fingers, knives, legs, arms, wings, ears’ (Genetti 1994: 68).¹⁴ Finally, the Thangmi numeral classifier <-gore> ‘houses, general things’ may be cognate with either the Dolakhā Newar classifier <-gar> ‘eggs, rice, rocks, noses, apples, balls, houses, stars, autos’ (Genetti 1994: 68) or the classifier <-gur> ‘(general classifier)’ (Genetti 1994: 69). The above examples provide powerful evidence of lexical similarities between the Sindhupālcok dialect of Thangmi and the Dolakhā dialect of Newar. Three further Thangmi numeral classifiers have no obvious cognates in Newar, and concomitantly, the five remaining classifiers present in Dolakhā Newar are not found in Thangmi.¹⁵ A full synchronic analysis of Thangmi numeral classifiers is offered in Section § 9 of Chapter 5.

¹³ I refer the reader to Austin Hale and Iswaranda Shresthacarya (1973) and Peri Bhaskararao and S. K. Joshi (1985).

¹⁴ According to Dörte Borchers, the Sunwar language (also known as Koĩnts) has a numeral classifier <-pa>, as in *nim-pa koel* (two-CLF leg) ‘two legs’ (personal communication). This may well be cognate with the Dolakhā Newar classifier <-pā> ‘fingers, knives, legs, arms, wings, ears’ described above.

¹⁵ There are only two numeral classifiers attested in the Dolakhā dialect of Thangmi: <-gore> for ‘non-human’ and <-ka> for ‘human’. While Thangmi <-gore> ‘non-human’ is likely to be

All Thangmi numeral classifiers are grammaticalised forms with no clear derivation from any related Thangmi nominal lexeme, apart from the classifier for human referents <-kapu>, which is also the Thangmi noun for ‘head’. The similarity between the Dolakhā Newar numeral classifiers and those found in the Sindhupālcok dialect of Thangmi are striking, and leaves open the question of whether these forms are borrowed or whether they reflect a close genetic relationship between the two languages.¹⁶ If the numeral classifiers are borrowed, the direction of the borrowing also remains unresolved. The Thangmi forms may be the more archaic as two of the Thangmi classifiers are disyllabic whereas their Newar counterpart are monosyllabic: Thangmi <-gore> and <-paṭe> versus Newar <-gar ~ -gur> and <-pta>. If the Thangmi classifiers were borrowed from Newar, then the suffixation of a velar nasal [ŋ] in the Thangmi classifier for ‘years’ <-daŋ> to the original Newar classifier <-da> ‘years’, and the presence of a trill [r] at the coda of the Thangmi classifier <-pur> are also difficult to explain.

Numeral classifiers are used to enumerate things in trade relations, and there is significant evidence of social and economic contact between the Thangmi and Newar groups, which may provide an argument for suggesting that the classifiers are indeed borrowed forms.¹⁷ Whether the shared classifiers can be used to argue for a close genetic relationship between the two languages or whether these impressive lexical similarities are merely a sign of intensive borrowing between Thangmi and Newar remains a central question.

4.2 Research on the Classical Newar language

Pursuing the idea of the alleged Thangmi-Newar link still further, I searched through lexical lists and dictionaries of contemporary and Classical Newar in search of possible correspondences. As this section illustrates, my findings add weight to the suggestion that when taken together, Newar, Thangmi and Barām form the higher-level grouping of Newaric.

cognate with the Dolakhā Newar classifier <-gur> ‘general classifier’, Thangmi <-ka> ‘human’ appears to have no cognates in Newar. The ‘human’ versus ‘non-human’ distinction is more reminiscent of Hayu, which has classifiers <-pu> for ‘human’ and <-uŋ> for ‘non-human’ (Michailovsky 1988: 123).

¹⁶ Quite why and how numeral classifiers attested in the Dolakhā dialect of Newar should have cognates in the Sindhupālcok dialect of Thangmi rather than the geographically closer Thangmi dialect spoken in Dolakhā remains unexplained. It is, however, possible that these classifiers were once also present in the Dolakhā dialect of Thangmi but are now no longer remembered, and are retained only in the Sindhupālcok dialect. This issue certainly warrants further exploration.

¹⁷ The cultural interdependence between the Newar of the Dolakhā and their Thangmi neighbours is dealt with in Section §9.2 of Chapter 2.

The time depth of the Classical Newar language has long been debated, as has its variational breadth. The controversy can be traced back at least as far as Hans Jørgensen, who described Classical Newar as simply ‘the language of the MSS [manuscript]’ (1936: 3). Five years later, in the *Preface* to his grammar of the language, he noted the explicitly ‘historical’ nature of Newar: ‘since the manuscripts...range from the fourteenth to the nineteenth century, and the natural changes in the language during this period have to some extent been reflected in them’ (1941: 3). The editors of the recently published *Dictionary of Classical Newari* are aware of the implications of this historical depth and make their position clear. In the *Introduction*, they state:

All we know at this stage is that Classical Newari is not a single homogenous monolithic stage nor a variety, dialect or stylistic label. (Malla *et al.* 2000: vii)

‘Classical Newari’ is an umbrella term used to describe the older forms of the language found in the 96 manuscript sources consulted for the dictionary, the chronological span of which ranges from 1115 A.D. to 1900 A.D. The editors reiterate their point by precluding comparisons between ‘Classical Newari’ and features of other so-called ‘Classical’ languages, such as Sanskrit, Greek, Arabic. As they see it, ‘Classical Newari’ is little more than a term of convenience used to separate a range of older Newar language varieties from Colloquial Newari (2000: viii).

This view is not shared by the Newar scholar Kashinath Tamot (*Kāśmāth Tamot*). Tamot believes that the existing linguistic divergences are more than ‘mere spatial variations—variations of individual dialects, (social/regional) or evidence of diaglossia [sic] (high style/low style)’ (Malla *et al.* 2000: viii). According to Tamot, there are ‘at least two stages of Classical Newari, i.e., Early and Late...This is approximately equivalent to the division of Nepal’s history into Early (879-1482) and Late (1482-1768) Medieval periods’ (2002: 13). Tamot is quick to point out the linguistic implications of this argument: Jørgensen’s dictionary would now only cover the Late Classical and Early Modern periods of the Newar language (from 1675 A.D. to 1859 A.D.). Tamot suggests that Early Classical Newar exhibited pre-Aryan features which were replaced by Sanskritic vocabulary in the Late Classical and Early Modern periods. Professor Kamal Prakash Malla (*Kamal Prakāś Malla*), Chief Editor and Project Leader of the *Nepal Bhasa Dictionary Committee*, is

palpably non-plussed by this theory and others of what he refers to as ‘Tamot’s hobby-horses’ (Malla *et al.* 2000: iv).¹⁸

At the 9th Seminar of the International Association for Tibetan Studies in June, 2000, Tamot presented a paper entitled ‘Some characteristics of the Tibeto-Burman stock of Early Classical Newari’.¹⁹ I was interested to find cognates between Thangmi as spoken to this day in Dolakhā and Sindhupālcok and certain ‘Early Classical Newar’ words, lexical items which were replaced by Sanskritic loans in Late Classical Newar. Subsequently, Tamot and I discovered that a number of Thangmi ritual words for body parts closely resemble Classical Newar forms, a discovery which lends further credence to the proposed closeness of the Newar and Thangmi peoples and their languages. In the following sections, I present an overview of the lexical similarities between Thangmi and Classical Newar.

4.3 Three classes of Thangmi and Classical Newar correspondences

Lexical items shared by Thangmi and Classical Newar fall into three categories. The first, and also the least spectacular, are words which are well-attested reflexes of Proto-Tibeto-Burman roots found across the genetically-related languages of Nepal and the higher Himalayas. That Thangmi and Classical Newar share these words does little more than reconfirm their membership in the Tibeto-Burman language family. The second class of shared items are Indo-Aryan loan words which have entered both Thangmi and Classical Newar. While many of the Tibeto-Burman languages of Nepal are considerably influenced by Indic, the Newar are the sole Tibeto-Burman people to have adopted both a Sanskrit literary tradition as well as a version of the Indo-Aryan caste system, a result of which is a heavily Sanskritised lexicon. The most likely explanation for these shared Indic loans is that one of the two languages loaned words from Sanskrit which were then, at a later date, borrowed by the other. Another possibility is that both Thangmi and Classical Newar were in contact with the same Indic language, perhaps at even approximately the same time. At any rate, as can be seen from the examples below, there are a number of shared Indic loans where one might have expected to find a non-loaned and native Tibeto-Burman form. The third and final class of lexical items shared by Thangmi and Classical Newar is by far the most interesting. This category consists of the numerous correspondences between the two languages, few (if any) of which are attested in other Tibeto-Burman languages spoken in the Himalayan region.

¹⁸ Sadly, this important lexicographical project on the Newar language was not without its tensions and disagreements. Malla writes of ‘unexpected and unhappy turns’ (2000: iii), which included the resignation of Kashinath Tamot, the Chief Compiler of the project.

¹⁹ Now published as an article with the same title, see Tamot (2002).

A brief disclaimer at this point would be prudent: Tibeto-Burman historical linguistics is still in its infancy in comparison with the depth of comparative and historical scholarship which exists for Indo-European languages.²⁰ It is likely that some of the lexical items I include in the proposed list of those shared by only Thangmi and Classical Newar will prove, over time, to be reflexes of Proto-Tibeto-Burman roots or cognate with elements found in other extant Himalayan languages.

The data are presented according to the three categories outlined above. Following each citation of a Proto-Tibeto-Burman form or Classical Newar word, its provenance is indicated by brackets with the following abbreviations: (B) for Benedict's *Sino-Tibetan*, (J) for Jørgensen's *Dictionary of the Classical Newarī*, and (NB) for the Nepal Bhasa Dictionary Committee's recent *Dictionary of Classical Newari*. Matisoff's recent *Handbook* builds on, and further develops, many of Benedict's early proto-forms, and I refer to these reconstructions in the forthcoming sections when cognates or reflexes are apparent.

4.3.1 *Shared common reflexes of Tibeto-Burman*

In this section, I present a list of Thangmi and Classical Newar words which are reflexes of well-attested Proto-Tibeto-Burman forms, or clearly cognate with lexical items in other extant Tibeto-Burman languages spoken in the Himalayas.

The reflexes of common Tibeto-Burman proto-forms include body parts, animals, food stuffs and verb roots. Reflexes of Proto-Tibeto-Burman *s-wa 'tooth' (B) are Thangmi *suwa* 'tooth' and Classical Newar *wā* 'tooth' (J); reflexes of Proto-Tibeto-Burman *kliy 'excrement' (B) are Thangmi *kli* 'excrement' and Classical Newar *khi* 'excrements' (J) and *khī* 'faeces' (NB); *(g-)yak 'armpit' (B) has reflexes *yakho* 'armpit' in Thangmi and *yāko* 'armpit' (J) in Classical Newar; Thangmi *nyu* ~ *nyu* ~ *nyī* 'brain' and Classical Newar *hni-pu* 'brain' (J) and *nhipu* 'brain' (NB) are cognate with Proto-Tibeto-Burman *nuk 'brain' (B); Proto-Tibeto-Burman *r-mi(y) 'man' has reflexes *mi* 'person, man' in Thangmi and *mim* 'man' in Classical Newar (NB); Proto-Tibeto-Burman *r-sa 'vein' (B) has reflexes *sasa* 'vein, tendon' in Thangmi and *śaśa* 'sinews, vein' in Classical Newar (NB); Proto-Tibeto-Burman *sya=śa 'meat' (B) has reflexes in both Thangmi and Newar indicating bovines, since these were once eaten by Newars (and still are eaten by the Thangmi): *sya* ~ *shya* 'cow' in Thangmi and *śā* 'cow' in Classical Newar (NB). The related forms *syaca* 'calf' in Thangmi and *sacā* 'calf' in Classical Newar (NB) are derived from Proto-Tibeto-Burman *sya=śa 'meat' (B) and *tsa 'child' (B) respectively. Proto-

²⁰ In a similar vein, James Matisoff writes of the 'present imperfect state of TB [Tibeto-Burman] historical phonology' (2000: 368).

Tibeto-Burman *s-rik=śrik ‘louse’ (B) has reflexes *sirik* ‘louse’ in Thangmi and *śi* ‘body louse’ in Classical Newar (NB); and Proto-Tibeto-Burman *lak ‘arm, hand’ (B) has reflexes *lak* ~ *la?* ‘hand, arm’ in Thangmi and *lā* ‘hand, arm’ (J) or *laka* ‘arm’ (NB) in Classical Newar.

Other reflexes for animal and organic words are as follows: Thangmi *amu* ‘eagle’ and Classical Newar *imā* ~ *yumā* ‘eagle’ (NB) from Proto-Tibeto-Burman *muw=məw ‘eagle’ (B); Thangmi *kucu* ‘dog’ and Classical Newar *khicā* ‘dog’ (NB) from Proto-Tibeto-Burman *kwiw=kəy ‘dog’ (B); Thangmi *kucuca* ‘puppy’ and Classical Newar *khicācā* ‘puppy’ (NB) from Proto-Tibeto-Burman *kwiw=kəy ‘dog’ and *tsa ‘child’ (B); and the related forms *ma-kucu* ‘bitch, female dog’ in Thangmi and *mā-khicā* ‘bitch’ in Classical Newar (J) from Proto-Tibeto-Burman *ma ‘mother’ and *kwiw=kəy ‘dog’ (B); Thangmi *naru* ‘horn’ and Classical Newar *na* ‘horn’ (J) and *na* ~ *nakura* ‘horn’ (NB), both containing a shared initial element alongside the reflex of Proto-Tibeto-Burman *krew=krəw or *ruŋ=rwaŋ ‘horn’ (B); Thangmi *naŋa* ‘fish’ and Classical Newar *nā* ‘fish’ (J) from Proto-Tibeto-Burman *ŋya ‘fish’ (B); Thangmi *pya* ‘pig’ and Classical Newar *phā* ‘hog, boar’ (J) or *phā* ‘pig, boar’ (NB) from Proto-Tibeto-Burman *pwak ‘pig’ (B); the Thangmi bound morph <sek> ‘fruit, round organic object’ and Classical Newar *se* ‘fruit, corn, grain’ (J) and *se* ‘fruits’ or *cākuse* ‘a kind of sweet yellow citrus fruit about the size of an orange’ (NB) from Proto-Tibeto-Burman *sey ‘fruit’ (B); Thangmi *jake* ‘rice’ and Classical Newar *jāke* ~ *ke* ‘rice, husked rice’ (NB) from Proto-Tibeto-Burman *dza ‘eat’ (B); Thangmi *chya* ‘salt’ and Classical Newar *chi* ‘salt’ (J) or *cī* ‘salt’ (NB) from Proto-Tibeto-Burman *tsa ‘salt’ (B); and Thangmi *marci* ‘chilli’ and Classical Newar *maracabhata* ‘chilli’ (NB) which are cognate with Yamphu *marchu* ‘Spanish pepper, red pepper, *Capsicum annum*’ (Rutgers 1998: 555) and similar terms in many other Tibeto-Burman languages, all probably derived from Sanskrit *maricaḥ* ‘pepper’.²¹ Two notable kinship terms are *nini* ‘husband’s sister, father’s sister’ (J) and *nini* ‘aunt, father’s sister’ (NB) in Classical Newar and *nini* ‘father’s sister’ in Thangmi, from Proto-Tibeto-Burman *ni(y) ‘aunt’ (B); and Thangmi *bubu* ‘elder brother’ and Classical Newar *phupa* ‘elder brother’ (NB) from Proto-Tibeto-Burman *puw=pəw ‘brother, older’ (B), another reflex of which is Kulung *bu* ~ *bubu* ‘elder brother’ (Tolsma 1999: 197).

Inanimate nouns with common reflexes are Thangmi *asku* ‘smoke’ and Classical Newar *kum* ‘smoke’ (NB) from Proto-Tibeto-Burman *kuw=kəw ‘smoke’ (B); Thangmi *asa* ‘oil’ and Classical Newar *so* ‘oil’ (NB) from Proto-Tibeto-Burman *sa-w ‘oil’ (B); Thangmi *uni* ‘day, sun’ and Classical Newar *hni* ‘day’ (J) and *nhī*

²¹ Chillies arrived in South Asia some time after the beginning of the 16th century.

‘day’ (NB) from Proto-Tibeto-Burman *niy=ŋəy ‘day’ (B); Thangmi *nasa* ‘soil, earth, ground’ and Classical Newar *cā* ‘soil’ (NB) are likely cognate with Tibetan *sa* ‘earth, the ground’ (Jäschke 1968: 568); Thangmi *rapa* ‘axe’ and Classical Newar *pāo* ~ *pā* ‘axe’ (NB) from Proto-Tibeto-Burman *r-wa=r-pwa ‘axe’ (B); Thangmi *khanou* ‘door, door-frame’ and Classical Newar *khā* ‘door’ (J) or *kāpā* ‘door’ (NB) from Proto-Tibeto-Burman *m-ka ‘door’ (B); Thangmi *ṭaye* ‘night’ and Classical Newar *cā* ‘night’ (NB) perhaps from Proto-Tibeto-Burman *ya ‘night’ (B); Thangmi *cinem* ‘iron’ and Classical Newar *ñam* ‘iron’ (NB) perhaps from Proto-Tibeto-Burman *syam=sam ‘iron’ (B); Thangmi *me* ‘fire’ and Classical Newar *mi* ~ *me* ‘fire’ (J) from Proto-Tibeto-Burman *mey ‘fire’ (B); Thangmi *me-thap* ‘fireplace’ and Classical Newar *mi-thap* ‘chimney’ (J) from the two Proto-Tibeto-Burman elements *mey ‘fire’ (B) and *tap ‘fireplace’ (B); Thangmi *kham* ‘word, tale, story’ and Classical Newar *kha* ‘word, tale, story’ (J) or *kham* ‘matter, fact, talk, dispute’ (NB) from Proto-Tibeto-Burman *ka ‘word, speech’ (B); Thangmi *siŋ* ‘tree, wood’ and Classical Newar *sima* ~ *siŋ* ‘tree, a plant, wood’ (NB) from Proto-Tibeto-Burman *siŋ ‘tree’ (B); and Thangmi *ulam* ‘path, road’ and Classical Newar *lam* ‘road, way, direction’ (J) or *lam* ‘way, road’ (NB) from Proto-Tibeto-Burman *lam ‘road, direction’ (B).

Common verb cognates and other grammatical particles are Thangmi *cabusa* ‘to carry’ and Classical Newar *ku buyu* ‘v.t., to carry’ (NB) from Proto-Tibeto-Burman *buw=bəw ‘carry on back or shoulders’ (B); Thangmi *ganḍu siŋ* ‘dry wood’ and Classical Newar *gaṃga siŋ* ‘dry wood’ (NB) from Proto-Tibeto-Burman *kaŋ ‘dry up’ (B); Thangmi *walḥa* ‘five’ and Classical Newar *ñam* ‘five’ (NB) from Proto-Tibeto-Burman *l-ŋa ~ b-ŋa ‘five’ (B); Thangmi *ca* ‘small, young, diminutive’, *caca* ‘very small’ and *cacha* ‘grandchild’ and Classical Newar *cā* ‘a young one (of animals)’ (J) or *cā* ‘child, young, small, diminutive suffix’, *cacā* ‘small, minor’ and *chaya* ‘grand-daughter, grand-son’ (NB) from Proto-Tibeto-Burman *tsa ‘child, grandchild, nephew, niece’ (B); Thangmi *pisa* ‘to give (away)’ and Classical Newar *pi-tě* ‘to give away’ (J) or *biye* ‘to give, to pay’ (NB) from Proto-Tibeto-Burman *biy=bəy ‘give’ (B); Thangmi *losa* ~ *lokxa* ‘to pour’ and Classical Newar *lu* ‘to pour’ (J) or *luya* ‘to pour’ (NB) from Proto-Tibeto-Burman *(m-)lu(w) ‘pour’ (B); Thangmi *lupsa* ‘to sink, to be submerged’ and Classical Newar *lop* ‘to sink, to be submerged’ (J) from Proto-Tibeto-Burman *lip and/or *nup~*nip ‘sink’ (B); Thangmi *sāisa* ‘to know’ and Classical Newar *saya* ‘to know, to understand, to be conversant with’ (J) or *sayā* ~ *saye* ‘to get notice, to know’ (NB) from Proto-Tibeto-Burman *syey ‘know’ (B); Thangmi *the* ‘self’ and Classical Newar *thao* ‘self’ (NB) from Proto-Tibeto-Burman *tay ‘self’ (B); Thangmi *duŋ* ‘inside’ and Classical Newar *duone* ‘inside’ (NB) from Proto-Tibeto-Burman *tsyu·ŋ=tu·ŋ ‘inside’ (B);

Thangmi *namsa* ‘to smell’ and Classical Newar *naṃṃa* ‘to smell’ (NB) from Proto-Tibeto-Burman *m-nam ‘smell’ (B); Thangmi *nuisa* ‘to laugh, smile’ and Classical Newar *nhira* ~ *nhile* ‘to laugh’ (NB) from Proto-Tibeto-Burman *m-nwi(y) ‘laugh’ (B); Thangmi *sisa* ‘to die’ and Classical Newar *sica* ‘to die’ (NB) from Proto-Tibeto-Burman *siy=səy ‘die’ (B); and Thangmi *su* ‘who?’ and *susu* ‘whoever?’ and Classical Newar *su* ‘who? (of persons only)’ (J) or *sū* ‘who’ and *susu* ‘whoever’ (NB) which are cognate with modern written Tibetan *su* ‘who?’ (Jäschke 1990: 573).

Certain Thangmi and Classical Newar words are also cognate with Sampang, a Kiranti language spoken in the north-eastern quadrant of Khoṭāñ district.²² Thangmi *chusa* ‘to fasten’ and Classical Newar *chuya* ‘to fasten, to attach’ (J) are cognate with Sampang *chuyma* ‘to fasten’; Thangmi *bok* ‘inflorescence of corn or rice flower’ and Classical Newar *bo* ‘flower’ (J) and *bo* ‘flower’ (Modern Newar *bun*) (NB) are cognate with Sampang *bun* ‘flower’; Thangmi *mesya* ‘buffalo’ and Classical Newar *mes* ‘buffalo’ (J) or *mesa* ‘buffalo’ (NB) are cognate with Sampang *mesi* ‘buffalo’ and Kulung *me:si* ‘water buffalo’ (Tolsma 1999: 220).

Overall then, the above examples only serve to demonstrate that Thangmi and Newar are Tibeto-Burman languages which contain reflexes of well-attested proto-forms and have cognates in extant Tibeto-Burman languages spoken across the Himalayas.

4.3.2 Shared Indo-Aryan loans

The second category presented here contains lexical items which both Thangmi and Classical Newar have borrowed from Indo-Aryan, and I have focussed solely on the loans which are particularly similar. Thangmi *aji* ‘mother-in-law’ and Classical Newar *aji* ‘grandmother (paternal and maternal)’ (J) or *aji* ‘grandmother’ (NB) may well have been loaned from Hindi *ājī* ‘paternal grandmother’ (McGregor 2002: 82); Thangmi *kapale* ‘forehead’ and Classical Newar *kapāra* ‘forehead’ (NB) are both cognate with Nepali *kapāl* ‘forehead, scalp’; Thangmi *kaṅṭu* ‘throat, neck’ and Classical Newar *kaṅṭhu* ‘throat’ (NB) are probably borrowed from a later reflex of Sanskrit *kaṅṭha* ‘throat, neck’; Thangmi *ṭupuri* ‘hat, cap’ and Classical Newar *tupuli* ‘a sort of head-gear’ (J) from Hindi or Maithili *ṭopī* ‘cap’; Thangmi *ṭhāi* ‘place, location’ and Classical Newar *ṭhāya* ‘place’ (NB) are likely etymologically related to Nepali *ṭhāū* ‘place’; Thangmi *dudu* ‘milk, woman’s breast’ and Classical Newar *dudu* ‘milk, the breast of a woman’ (J & NB) may be loaned from Nepali (or another neighbouring Indo-Aryan language) *dud* or *dudh* ‘milk, female breast, udder’; Thangmi *ḍumla* ‘common fig, *Ficus carica*’ and Classical Newar *dubala* ‘*Ficus*

²² The Sampang data are provided by René Huysmans, via personal communication.

racemosa (formerly known as *Ficus glomerata*) may derive from Sanskrit *uḍumbara* ‘*Ficus racemosa*’; Thangmi *sakalei* ‘all, everything, everyone’ and Classical Newar *śakala* ‘everybody’, *sakale* ‘all’ (NB) and *sakale* ‘all’ (J) likely derive from a later reflex of Sanskrit *sakala* ‘whole’; Thangmi *paṭasi* ‘women’s traditional dress’ and Classical Newar *patāsi* ‘the lower garment’ (J) or *patāse* ‘a woman’s lower garment’ (NB) may derive from a later reflex of Sanskrit *paṭaḥ* ‘cloth’ or Nepali *pāṭ* ‘flax, fibre’; Thangmi *naka* ‘new’ and Classical Newar *naka* ‘new’ (NB) may derive from a later reflex of Sanskrit *nava* ‘new, fresh’; Thangmi *sewa* ‘greetings, hello’ and Classical Newar *sevā* ‘a term of address to show respect to elders’ (NB) are derived from a later reflex of Sanskrit *sevā* ‘attendance (upon someone), servitude’; and finally Thangmi *makar* ‘monkey’ and Classical Newar *markaṭ* ‘monkey’ (J) or *mākarha* ‘monkey’ (NB) are most probably loaned from Nepali *markaṭ* ‘monkey’ and ultimately derive from Sanskrit *markaṭa* ‘monkey’.

As outlined above, Newar has a highly Sanskritised lexicon and it is no surprise that even words which might be considered part of the core lexicon, such as ‘very’, ‘milk’ or ‘breast’, have been loaned from Indo-Aryan. Surprising, however, is that Thangmi has also borrowed these terms, and furthermore, that the loans seem to have undergone similar phonological shifts in both languages. Examples are the reduplicative *dudu* ‘milk’ from Indo-Aryan *dud* or *dudh*, and the extra syllable added to the loan for ‘hat, cap’ as in Thangmi *tupuri* and Classical Newar *tupuli*, from Indo-Aryan *ṭopī*.²³

The most plausible explanation for this similarity in loaned words is that one of the two languages borrowed words from a neighbouring Indic language which were then at a later date borrowed ‘once-removed’ into the second language. The sequence of these loans was most probably Classical Newar borrowing from Indo-Aryan and then Thangmi borrowing an Indic or Sanskritised lexical item from Newar. Due to the high level of literacy and the extensive written tradition of Newar civilisation, loans directly from Sanskrit into Classical Newar were commonplace. For Thangmi, however, which remains to this day an unwritten language spoken far from any urban centre of learning, direct loans from Sanskrit are less likely. The transfer scenario outlined above would support the hypothesis that the Thangmi and Newar languages (and hence their speakers) were in close contact with one another from an early date. In the absence of such early contact, one would have expected

²³ Although less phonologically persuasive, other possible shared Indo-Aryan loans may be Thangmi *athu* ‘joint of the body’ and Classical Newar *athi* or *āthi* ‘joints, articulations’ (NB), both perhaps from a later reflex of Sanskrit *asthi* ‘bone, joint’ or Hindi *asthi* ‘bone’ (McGregor 2002: 70); and Thangmi *athe* ‘very’ and Classical Newar *ati* ‘very, exceedingly’ (J) which may have been loaned from Maithili, Nepali or Hindi *ati* ‘very, very much’.

Thangmi to borrow lexical items directly from Nepali (rather than from Sanskrit by way of Newar) when the Nepali language was brought to Dolakhā and Sindhupālcok by Nepali-speaking Indo-Aryan settlers.

4.3.3 *Lexical correspondences specific to Thangmi and Newar*

The final category, which is the most interesting one, comprises those lexical similarities which I believe to be shared by only Thangmi and Classical Newar and which are probably not cognate with other Tibeto-Burman languages. Although recent articles and conference papers have refined and added to Benedict's list of Tibeto-Burman reconstructions, for reasons of space they are not included here. I have resisted the temptation to order the lexical similarities into classes (such as animate nouns, body parts and verbs) since this would impose a further arbitrary hierarchical order on the data. As far as possible, I have followed the alphabetical order of the Newar dictionaries, thus facilitating cross-referencing for those interested readers.

Thangmi *elepe* 'spleen' is cognate with Classical Newar *al-pe* 'spleen' (J) and *alape* 'spleen' (NB), for which Matisoff reconstructs *r-pay 'spleen' (2003: 208); Thangmi *pin* 'fingernail' may be cognate with Classical Newar *r* 'nail' (NB); Thangmi *ekaṭe* 'alone' is cognate with Classical Newar *ekāṭa* 'alone' (NB), the first syllable of both being cognate with and derived from Sanskrit *eka* 'one, a, only, alone, single'; Thangmi *kaṭasa* 'to quarrel' closely resembles Classical Newar *kacāda* ~ *kacāda* 'quarrel, dispute' (NB); Thangmi *kapale kosa* 'skull' similarly resembles Classical Newar *kapāla kosa* 'skull' (NB); Thangmi *kasyu* 'boil, pimple' and Classical Newar *kasu* 'boils' (NB) or Classical Newar *cāsu kacha* 'a pimple that itches' (NB) are most probably cognate, as are Thangmi *kimi* 'tape worm' and Classical Newar *kimi* 'hook worm' (NB).

Other plausible lexical correspondences include Thangmi *cyuku* 'ant' and Classical Newar *kumicā* 'white ant, termite' (NB); Thangmi *kosa* 'bone' and Classical Newar *kvase* ~ *kosa* 'bones' (NB); Thangmi *papasek* 'testicles' and Classical Newar *si-pā* 'the testicles' (J) or *kvāse* 'testicles' (NB), for which Matisoff reconstructs *səw 'testicles, virility' (2003: 182); Thangmi *ukhiṅ* 'dark' and Classical Newar *khiṅ* 'dark, darkness' (J) or *khiṅni* 'dark' (NB); Thangmi *gui* ~ *gwi* 'thief' and Classical Newar *khu* 'thief' (NB); and Thangmi *khen* 'face' and Classical Newar *khem* 'face' (NB). The Thangmi individuating suffix *guri* may be cognate with Classical Newar *guri* 'a classifier denoting place' (NB); Thangmi *naṭe* 'cheek' resembles Classical Newar *natāl* 'cheek' (J) and Modern Newar *nyatāḥ* 'cheek' (NB); Thangmi *ṭakadu* 'sweet' is most likely cognate with Classical Newar *cāku* 'sweet' (J) and *cāku* 'sweet thing, molasses' (NB); and Thangmi *cime* 'hair (on the

scalp)’ is most likely cognate with Classical Newar *cimī* ‘the hair (of the body)’ (J) and *cimilisāṃ* ‘body hair’ (NB), for which Matisoff reconstructs *mil ~ *mul ~ *myal ‘hair (body)’ (2003: 602).

Other possible cognates are Thangmi *jakcho* ‘wheat’ and Classical Newar *cho* ‘wheat’ (J) or *co* ~ *cho* ‘wheat’ (NB); the Thangmi affable suffix *che* and Classical Newar *che* ‘2.s. you (used mostly in addressing superiors or equals)’ (J) or *cha* ‘you’ (NB); Thangmi *cacha jyamari* ‘granddaughter’s husband’ and Classical Newar *chaya jīri* ‘granddaughter’s husband’ (NB); the Thangmi verb *chyosa* ~ *thosa* ‘to send’ and Classical Newar *choya* ‘to send, to dispatch’ (J) or *choye* ~ *choya* ‘to send’ (NB); Thangmi *jukun* ‘only’ and Classical Newar *jak* ‘only’ (J) or *juko* ‘only’ (NB); the Thangmi noun *jet* ‘work’ and Classical Newar *jyā* ‘work’ (NB); and Thangmi *jyanḡaṅeṅ* ‘bird’ and Classical Newar *jhaṅgara* ‘a bird’ (NB). Another set of lexical similarities shared by the two languages includes Thangmi *thumsa* ‘to bury’ and Classical Newar *thumñā* ~ *thumne* ‘to bury’ (NB); the Thangmi verb *themsā* ‘to destroy, to break down’ and Classical Newar *thoña* ~ *thone* ‘to demolish, to destroy’ (NB); Thangmi *daṅ* ‘year’ and Classical Newar *ḡa* ~ *daṅ* ‘year’ and Modern Newar *daṅ* ‘year’ (NB); the Thangmi male clan *ḡaṅguri* and Classical Newar *ḡhaṅguri* ‘a Newar caste’ (NB); the Thangmi kinship term *tete* ‘elder sister’ and Classical Newar *tatā* ‘an elder sister’ (J) or *tatāju* ‘elder sister (hon.)’ (NB); Thangmi *thope* ‘broom, sweep’ and Classical Newar *tuphe* ‘a broom’ (J) or *tuphi* ‘broomstick, brush’ (NB); Thangmi *du* ‘tiger, leopard, wild cat’ and Classical Newar *dhu* ‘tiger’ (J) or *tedu* ‘leopard’ (NB); the Thangmi verb *thisa* ‘to touch’ and Classical Newar *thiye* ‘to touch’ (NB); Thangmi *thumsa* ‘to immerse’ and Classical Newar *thune* ‘to immerse’ (NB); the Thangmi shamanic and ritual ethnonym for themselves *thani* and Classical Newar *thāni* ‘one kind of caste’ (NB);²⁴ and the Thangmi noun *toṅ* ‘home-made beer’ and Classical Newar *thvaṅ* ‘beer’ (NB).

Further Thangmi-Classical Newar lexical correspondences include: Thangmi *duṅ bisa* ‘to enter (inside)’ and Classical Newar *duṅbiya* ‘to enter, to offer’ (NB); Thangmi *dudu pur* ‘nipple of the breast’ and Classical Newar *dudu pipīri* ‘nipple of the breast’ (NB); Thangmi *nama* ‘with’ and Classical Newar *na* ‘with’ (NB); the Thangmi plural suffix *pali* and Classical Newar *paṅi* ‘plural suffix’ (NB); Thangmi *paṅ* ‘sour’ and Classical Newar *pānu* ‘sour’ (NB); the Thangmi transitive verb *palsa* ‘to chop’ and Classical Newar *pāle* ‘to cut, to behead’ (NB); Thangmi *priṅ* ‘outside’ and Classical Newar *pi* ~ *piṅ* ‘outside’ (NB); the Thangmi transitive verb *busa* ‘to cover, fill’ and Classical Newar *puya* ‘to cover, to fill’ (NB), for which Matisoff reconstructs *pun ‘wrap, cover, wear’ (2003: 495); the Thangmi noun *puya*

²⁴ As intriguing as this Newar definition is, no further information is provided.

‘seed, seedling’ and Classical Newar *pu* ‘seed’ (J) or *pū* ‘seed’ and *puvā* ‘paddy seedlings’ (NB); the related Thangmi form *puyapasa* ‘grains and seeds’ and Classical Newar *puvāpāsā* ‘grains and seeds’ (NB); and the Thangmi kinship term *pāiri* ‘elder brother’s wife’ and Classical Newar *pairabe* ‘elder brother’s wife’ (NB).

Other possible lexical correspondences are Thangmi *pokole* ‘knee’ and Classical Newar *paulṭ* ‘knee’ (NB); Thangmi *phaṭu* ‘pumpkin’ and Classical Newar *phat-si* ‘a kind of pumpkin’ (J) or *phatase* ~ *phatse* ‘pumpkin’ (NB);²⁵ Thangmi *phasa* ‘wind, storm, air’ and Classical Newar *phas* ‘air, wind’ (J) or *phasa* ‘wind’ (NB); Thangmi *pebu* ‘field’ and Classical Newar *bu* ‘a field’ (J) or *bū* ‘a field’ (NB); Thangmi *bosa* ‘to grow’ and Classical Newar *boye* ‘to grow, to come up’ (NB); the Thangmi verb *mraṅsa* ‘to swell’ and Classical Newar *maṅ-gwo* ‘swelling’ (J) or *māna* ~ *māne* ‘to swell’ (NB); the Thangmi noun *maṅa* ~ *maṅiṅ* ‘bread’ and Classical Newar *mādhe* ‘bread’ (NB); Thangmi *mesyaca* ‘buffalo calf’ and Classical Newar *mesacā* ‘buffalo calf’ (NB) (a composite form particular to Thangmi and Newar, although the constituent elements are well-attested throughout Tibeto-Burman); Thangmi *moṭe* ‘soybean’ and Classical Newar *mvāca* ‘soybean’ (NB); and the Thangmi transitive verb *rasa* ‘to bring’ and Classical Newar *rāsa* ~ *rāye* ‘to seize, catch’ (NB).

A particularly interesting lexical similarity shared by Thangmi and Classical Newar is Thangmi *libi* ‘after, later, behind’ and Classical Newar *lithe* ‘later’, *li* ‘after’ and *livā* ‘afterwards’ (NB). Further correspondences are Thangmi *lukusa* ‘back, backbone’ and Classical Newar *luku* ‘back of body’ (NB); Thangmi *khaśu* ‘cloud’ and Classical Newar *śu* ‘cloud’ (NB); and Thangmi *sumaka* ‘quietly’ and Classical Newar *sumhaka* ‘quietly’ (NB). The final list of correspondences are those lexical items found only in Jørgensen’s *Dictionary* and not present in the newer *Dictionary of Classical Newari*: Thangmi *wasa* ‘to plough’ and Classical Newar *wāsā* ‘a plough’; Thangmi *aṅaldu* ‘ashamed’ and Classical Newar *nālā-pu* ‘ashamed, shame’; Thangmi *baṭi* ‘cat’ and Classical Newar *bhaṭi* ‘a cat’; Thangmi *makarpapa* ‘spider’ and Classical Newar *mā-khā-pi-khā* ‘a spider’; Thangmi *maṅ* ‘body’ and Classical Newar *hma* ‘a body’; Thangmi *laṅga* ‘courtyard’ and Classical Newar *lam-hñe* ‘a yard, a court’; Thangmi *sebi* ‘leather, hide, skin’ and Classical Newar *se-bu* ‘leather’; and Thangmi *ṭamsil* ‘marrow’ and Classical Newar *sel* ‘marrow’.

5. Concluding thoughts on the genetic affinity of Thangmi

Section §4.3.3 above contains over seventy likely cognates between Thangmi and Classical Newar, many of which may ultimately turn out to be derived from Proto-

²⁵ Both may be derived from Nepali *pharsi* ‘pumpkin’.

Tibeto-Burman roots, but which, at any rate, appear to have undergone a shared history in some earlier stage of Thangmi and Newar. The number of reflexes of Tibeto-Burman proto-forms may actually be quite high, but good reconstructions are hard to come by, and I can only hope that colleagues may lend a hand in weeding out those lexical items which are found in other Tibeto-Burman languages. However, even if half of the above proposed lexical similarities between Thangmi and Classical Newar turn out to be reconstructible to Proto-Tibeto-Burman, over thirty-five specifically shared lexical similarities remain. As mentioned at the outset, Shafer's argument for Thangmi and Barām relatedness was based on nine lexical similarities shared by the two languages, seven of which may now be discounted as they are widely attested in other Tibeto-Burman languages. Even though only two of Shafer's proposed similarities remain, his hunch of a Thangmi-Barām link has been corroborated by more recent research by van Driem and myself. While many Tibeto-Burman languages of Nepal have some lexical cognates with either Thangmi or Classical Newar, to my knowledge there is no other language which shares as many lexical correspondences with Thangmi and Classical Newar as these two languages share with one another.

I conclude as I started, by asking a question. Should the similarity between Thangmi and Classical Newar simply be put down to borrowing, or does it reflect a deeper genetic relationship? If we opt for the more cautious explanation, putting the similarities down to cultural contact and lexical borrowing, then the question remains as to what type of early contact situation existed in which the speakers of these two languages could have exchanged so much so long ago.²⁶ If, on the other hand, we choose to conclude that the lexical similarities shown above are an indication of a close genetic relationship between Thangmi and Newar, then sound evidence from the fields of historical phonology and comparative morphology must be produced to support this suggestion.²⁷

²⁶ Tej Ratna Kansakar (*Tej Ratna Kansakār*), a leading scholar of the Newar language and Tibeto-Burman linguistics, is unconvinced by the argument for a close genetic relationship between Thangmi and Newar. He suggests that the linguistic and cultural links between the two groups are most likely the result of 'contact-induced changes' and that there is historical evidence to show that the Newar, wherever they settled, sought the assistance of 'various caste groups to fulfil religious, social and ritual functions' for them. Other than the Thangmi of Dolakhā, a further example Kansakar offers is of Tibetans in their native Lhasa, who were conscripted to play a ritual role in Newar festivals (personal communication, 18 September, 2000).

²⁷ To quote van Driem, the 'current impression is that the older the Newar vocabulary, the more specific lexical correspondences can be identified with Thangmi and Barām' (2001: 761).

It will be interesting to examine further linguistic evidence from the Dolakhā dialect of Newar when it becomes available. Genetti has suggested that many of the Thangmi lexical items presented here have clear cognates with Dolakhā Newar (personal communication, February 2001), which is to be expected given the socio-cultural links between the two groups outlined in Section §9.2 of Chapter 2. Genetti writes of Dolakhā as a ‘centre for trade and commerce’ (1994: 8), but also of the ‘relative isolation of Dolakhā as compared to Kathmandu’ (1994: 8). It is precisely this isolation that van Driem sees as crucial in determining the relative antiquity of the Dolakhā dialect of Newar:

the original Newar grammatical system remains more intact in the language of the descendants of the early Newar mercantile colonists in Dolakhā than in the innovative prestige dialects spoken in Kathmandu and Pāṭan. (2001: 766)

On account of the geographical location of the town of Dolakhā, Genetti suggests that the ‘Dolakha people would have had more contact with the Kiranti peoples of the east’ (1994: 8). In light of the data presented above on the verbal agreement morphology shared by Kiranti languages and Thangmi on the one hand, and the lexical correspondences between Thangmi and Newar on the other, Genetti’s proposal is particularly interesting. While Genetti dates the split between the Kathmandu and Dolakhā dialects of Newar to a ‘minimum of seven hundred years ago, and possibly much longer’ (1994: 8), van Driem suggests that the ‘divergence between the Kathmandu Valley dialects and Dolakhā Newar may perhaps be datable to a period of unrest between 750 and 983 A.D.’ (2001: 766), leading to a stability of the linguistic community which in turn contributed to the ‘evident archaism of Dolakhā verbal morphology’ (2001: 766).

In order to take the study to a deeper level beyond the inspection and comparison of surface forms, the next step in the analysis of the Thangmi-Newar link will be to determine whether there are any phonological correspondences between the two languages. Only then will we learn more about the essence of the relationship between Thangmi and Newar, and the relative position of both languages in the *Stammbaum* of Tibeto-Burman.