

Case Compounding in the Bodic Languages

Michael Noonan

University of Wisconsin-Milwaukee

1. INTRODUCTION

Case compounding has received a certain amount of attention in the literature in recent years, in particular the phenomenon known as *Suffixaufnahme* [e.g. in Plank 1995a] and the various sorts of case compounding in Australian languages, some of which manifest *Suffixaufnahme* and other types of case compounding. There has been relatively less attention paid to the phenomenon outside of these two areas of research, and this paper is an attempt to rectify the situation somewhat by presenting data from a large family of languages, the Bodic languages, spoken in the areas straddling the spine of the Himalayas.

1.1 BODIC LANGUAGES: The Bodic languages are a large and ramified branch of the Tibeto-Burman languages.¹ It is a controversial grouping in the sense that there is no consensus as to what should be placed within it and indeed whether it is a legitimate genetic grouping at all or just an assemblage of Tibeto-Burman languages that have been spoken in the region of the Himalayas long enough for the languages to have influenced each other in a variety of ways. A figure illustrating possible relationships among the Bodic languages can be found in Appendix 1.

For our purposes, nothing crucial hinges on the assumption that these languages form a genetic grouping. It suffices that structurally these languages share certain features, among which is a tendency to compound markers of case within a phonological word.

1.2 DEFINITION OF CASE COMPOUNDING: We will define case compounding as the inclusion of two or more case markers within a phonological word. In this paper, we are concerned exclusively with case compounding within nominal words, which is to say that we are not concerned with instances of case compounding within verbal words or in those of adjectives, though we will consider case compounding in nouns used adnominally. Further, we are concerned with case only in two of its functional domains: *relational* case, which relates core arguments to the verb and peripheral arguments to the clause; and *adnominal* case, which relates one NP to another. We are not concerned with case usage as it relates specifically to subordinate verbals either with regard to the case marking of whole clauses or to arguments within these clauses headed by subordinate verbals; we are also not concerned with *modal* case (Evans 1995), whereby a case marker with locative relational function may, under certain circumstances, code clausal tense, mood, and aspect on certain non-subjects.

¹ The work reported on in this paper has been supported by the following grants from the National Science Foundation: DBC-9121114, SBR-9600717, and SBR-9728369.

1.3 KINDS OF CASE COMPOUNDING: Given the definition of case compounding above, we can distinguish at least the following sorts of case compounding crosslinguistically.²

a. case stacking: This sort of case compounding is the functional equivalent of preposition compounding in languages like English, e.g. *into the house*, *onto the chair*, etc. With this type, two independently occurring case affixes are used together to describe a complex trajectory, one that is understood as a combination of the meanings of the two case affixes. The Tamangic language Chantyal provides an example (Noonan 2003a, 2005):

- (1) tʰim-nʰari-gəmsə
 house-INES-ABL
 'out from inside the house'

b. derivational: This is Austin's (1995) term for a situation whereby one case serves as the 'basis' [or 'founding form'] for another, which is not found independently without the first.³ The Tocharian A declension of 'teacher' can serve as an example:

(2)

	SINGULAR	PLURAL
NOMINATIVE	kæʃʃi	kæʃʃiŋ
ACCUSATIVE	kæʃʃim	kæʃʃis
GENITIVE	kæʃʃiyāp	kæʃʃiffi
INSTRUMENTAL	kæʃʃinyo	kæʃʃisyo
PERLATIVE	kæʃʃinā	kæʃʃisā
COMITATIVE	kæʃʃinaffæl	kæʃʃisaffæl
ALLATIVE	kæʃʃinatʃ	kæʃʃisatʃ
ABLATIVE	kæʃʃinæʃ	kæʃʃisæʃ
LOCATIVE	kæʃʃinaŋ	kæʃʃisaŋ

The genitive singular and nominative apart, the other cases are based on the accusative [historically *-in* and *-is* for the singular and plural, respectively] to which postpositions which once had governed the accusative fused. As markers of case, these former postpositions do not occur without the historic accusative.

c. referential: This term is also from Austin (1995) and, in the context of case compounding as here defined, refers to a set of locative nominals which can be interpreted either adnominally [as in (3)] or adverbially [as in (4)]. For either sort, these

² The terminological framework used here draws from Austin (1995), Plank (1995b), and Moravcsik (1995). I exclude from the typology that follows instances where two forms signaling the same case are found in a single word without substantially affecting the meaning. For example, Huber (2005:61) observes that in Kyirong Tibetan the ergative may be doubled as a way of clarifying or emphasizing the ergative marking which might otherwise not be clear. So **pu** 'son, boy' has a regular ergative **py**, but in addition there is a reinforced ergative **py-ge** containing **-ge**, an ergative marker ordinarily suffixed to stems with final consonants or long vowels.

³ Austin's definition of derivational case compounding is broader than mine and excludes the last limiting clause; it could be interpreted to include at least some instances of what I call here case stacking. His definition is, no doubt, appropriate for the Australian languages he was concerned with, but mine is intended to deal with a larger range of data.

are marked with one case indicating location or direction and another referencing the NP it modifies or refers to. The first example is from Warlpiri (Hale 1982),

- (3) ṅarrka-ṅku ka yankirri luwa-rni ṅapa-ṅka-rlu
 man-ERG PRES emu shoot-NONPAST water-LOC-ERG
 'The man at the waterhole is shooting the emu.'

and the second is from Jiwarli (Austin 1995):

- (4) thuthu-ngku juma-rti-nha yanga-rninyja warlpari-lu
 dog-ERG child-PL-ACC chase-PAST south.ALL-ERG
 'The dog chased the children south.'

The referential type is clearly related to the adnominal types described below.

d. headed adnominal, with a case-marked head: This is the phenomenon described in detail in Plank (1995a) and called there *Suffixaufnahme*. It involves the agreement by a nominal modifier of the case of its head while retaining its own adnominal case, typically a genitive, as in the following Anguthimri example (Schweiger 1995):

- (5) waṭayi-yaṃṛa-ma pat'a-ma
 old.man-GEN-ABL canoe-ABL
 'from the old man's canoe'

e. headed adnominal, with no case marking on the head: This phenomenon, called *Suffixhäufung* in Plank (1995a), is similar to the prototypical cases of *Suffixaufnahme* except that the head is not marked for case, while the modifier is marked both for the adnominal case and the case for the case of the entire NP. An example from Arinda follows, reported in Moravcsik (1995):

- (6) worra ingkata-kana-la
 son chief-GEN-ERG
 'the chief's son'

In this example, the ergative is understood to be associated with the NP defined by the head, *son*.

f. simple headless adnominal: This construction is referred to as *hypostasis* in Plank (1995a). In this construction a headless possessor is marked with the genitive [or other adnominal case] and also with the case indicating the grammatical function of the whole, headless NP within the clause. An example follows from Chantyal:

- (7) nə-ye-sə kḥi-ye-ra jḥi-i
 1S-GEN-ERG 2S-GEN-DAT hit-PERF
 'Mine bit yours'

In this example, the ergative and dative cases are assigned to NPs filling subject and direct object slots, respectively. The genitive-marked NPs filling these slots take on the clausal case as well.

g. complex attributive nominal: In this construction, a case-marked noun is further marked with a nominalizer-attributive affix,⁴ and the resulting noun may be further case-marked as appropriate for its nominal role. Consider the following examples from Chantyal:

⁴ The syncretism of nominalization and attribution is typical of the Bodic languages (Noonan 1997, 2007).

- (8) a. məŋgəle-ri-wa mənchi
 Mangale-LOC-NOM person
 'person from Mangale'
 b. məŋgəle-ri-wa kha-i
 Mangale-LOC-NOM come-PERF
 'someone/something from Mangale came'
 c. məŋgəle-ri-wa-ma
 Mangale-LOC-NOM-PL
 'those from Mangale'
 d. məŋgəle-ri-wa-ma-sə na-ra mara-i
 Mangale-LOC-NOM-PL 1S-DAT see-PERF
 'those from Mangale saw me'

The adnominal in (8a) is interpreted as a noun and fills a clausal noun slot in (8b). It is pluralized in (8c), and in (8d) acquires the ergative case appropriate for its grammatical role within the clause. Case marked forms like these may acquire an additional nominalizer-attributive affix, as in (9):

- (9) məŋgəle-ri-wa-ma-siŋ-wa photo
 Mangale-LOC-NOM-PL-COM-NOM photo
 'the photo belonging to the people from Mangale'

This, in turn, may acquire an appropriate case marker indicating the grammatical status of the whole, as in (10):

- (10) na-sə məŋgəle-ri-wa-ma-siŋ-wa-ra dekhə-i
 1S-ERG Mangale-LOC-NOM-PL-COM-NOM-DAT show-PERF
 'I showed it to the owners from Mangale'

Nominalization in this sense is a recursive process in Chantyal and other languages employing this construction and may involve case compounding under the definition given here. **məŋgəle-ri-wa-ma-siŋ-wa-ra** contains three case clitics: a locative, a comitative, and a dative, all within the same phonological word.⁵

2. CASE COMPOUNDING IN THE BODIC LANGUAGES

2.1 ATTESTED TYPES IN BODIC: Of the seven types of nominal case compounding identified in the last section, the Bodic languages attest four:

- (11) a. case stacking
 b. derivational
 f. simple headless adnominal
 g. complex attributive nominal

⁵ Edith Moravcsik (p.c.) has suggested that the classification of types of case compounding provided in this section could be reduced to a smaller number of distinct types: A. *Case compounding on heads*, which would include as subtypes a. *case stacking* and b. *derivational case compounding*; and B. *Case compounding on modifiers*, which would include as subtypes d. *Headed adnominal, with a case-marked head*, e. *Headed adnominal, with no case marking on the head*, and, as a single subtype c/f/g. which would be retitled *Case marker on modifier, no head present*. The classification given in §1.3 is preserved here in part because it serves the narrow purposes of this paper which is to present a discussion of case compounding in the Bodic languages.

Case stacking is quite widely attested, at least historically, in most branches of the family. Derivational case compounding is not especially common, but it is certainly attested. The simple headless adnominal type is not frequently mentioned in grammatical descriptions, so its frequency is difficult to assess; however, an informal sampling seems to indicate that it is probably very common in this group. The complex attributive nominal type is probably not common, but most descriptions of Bodic languages are not sufficiently detailed so as to make a clear assessment of the frequency of this construction possible. Each of these types will be discussed in the sections below.

2.2 CASE IN BODIC:

The Bodic languages are strongly verb-final and agglutinating. There are no concord classes, though in a few cases adjectives borrowed from Indic languages may reflect 'natural' gender, displaying a special form when modifying nouns designating human females.

With regard to case, within the nominal word the order is STEM+NUMBER+CASE, though in a few uncommon instances, number and case may be marked with a fused form. Almost all Bodic languages are morphologically split-ergative, though the semantic-pragmatic basis for the split varies considerably within the group. A few languages, *e.g.* some Tamangic languages such as Chantyal and Gurung, are fairly consistently ergative. As demonstrated by Bickel (2003), some languages [*e.g.* Belhare] may exhibit both A/P and S/P pivots; other languages [*e.g.* Chantyal] effectively have no pivots. Primary object marking (Dryer 1986) is widespread in the group: the case that marks indirect objects [usually called 'dative'] also marks high animacy direct objects. In all languages an unmarked [or zero-marked] noun is used for at least some direct objects and typically for other nouns as well: such nouns are usually referred to as 'absolutives'.

Postpositions almost always govern either the genitive or an absolute. Instances of postpositional governance of other cases are very rare. As an example of this phenomenon in Bodic, Koshal (1979) reports that in Ladakhi the postposition **kə** 'on, above' governs the ablative and the postposition **phi** 'for' governs the dative, but it would be interesting to investigate the matter to determine if this is truly an instance of governed case or an instance of case stacking since cognates of both supposed postpositions occur as case clitics in closely related languages.

2.2.1 CASE EXPRESSED BY CLITICS: Though the literature on Bodic languages generally refers to case markers as 'suffixes' and occasionally as 'postpositions', the 'core' case markers [those coding grammatical cases and the basic local ones] are usually best analyzed as clitics: they are not postpositions because they are not prosodically independent nor may they occur in free discourse independently from the nouns to which they are bound; they are not suffixes because when NPs are coordinated, the case forms are ordinarily not repeated for each NP, and where word order in particular languages permits, the case forms can attach to words other than the nouns to which they are logically associated since these clitics are always phrase-final clitics. As an illustration of

the phrase-final enclitic status of the case markers in a typical Bodic language, consider the following data from the Tamangic language Nar-Phu:

- (12) ɲâ tile [čhupruŋ nartwɛ-re] ni-čín
 1S yesterday Nar Phu-DAT.LOC go-PAST
 'I went to Nar and Phu yesterday'
- (13) paŋčɛn-se [nôkyu thɛpɛ mɸilaŋ-re] šê-čín
 leopard-ERG dog big black-DAT.LOC kill-PAST
 'The leopard killed the big, black dog'

The dative-locative case marker **-re** illustrated in the sentences above is prosodically part of the word to which it is attached: in Nar-Phu, for example, the tone associated with the noun or adjective spreads onto the case clitic **-re** [Noonan 2003b]. In (12), the conjoined locative NPs share the dative-locative case clitic. In (13), the direct object NP is marked with the dative-locative clitic [Nar-Phu is a primary object language in the sense of Dryer 1986]: note that the clitic is attached to the final word in the NP, not to the head.

There are generally no declensional classes and, pronouns apart, few paradigmatic irregularities. Where they occur, their grammatical import, if not necessarily their historical origin, is usually fairly transparent. For example, in Kathmandu Newari the plural affix is **-to** for ordinary animate nouns, but **-pī** for honored referents: cf **khica-to** 'dog's' and **pasa-pī** 'friends' [Hargreaves 2003]. True irregularities may have more complex origins. In the Tamangic language Chantyal, five high frequency nouns have an irregular locative, not otherwise attested in Tamangic nouns, though it is found in various directionals:⁶

- (14) them-əŋ 'house, home'
 kɸyam-əŋ 'path, way'
 gāw-əŋ 'village'
 Bini-əŋ 'Beni'
 kyels-əŋ 'field'

[The locative is otherwise rendered by either **-ri** or **-ra**, depending on certain predictable semantic factors (Noonan 2005).] Of these five nouns, two ['village' and 'Beni'] are of Nepali origin, and so cannot instantiate an old, retained pattern; of the other three words, one demonstrates an otherwise unattested phonological irregularity: 'house' is otherwise **thim**. The cognates of these words in the other Tamangic languages do not exhibit any sort of irregularity. These and other facts suggest that these forms were either borrowed from or influenced in their development by the Bodic language Magar, whose generic locative is **-aŋ**: much of the Chantyal population is descended from ethnic Magars (Noonan 1996, 2008b).

⁶ By directionals, I mean words with meanings like 'down [static]', 'downward', 'up [static]', 'upward', etc. In these words, the element **-ŋ** indicates static location. With the nouns in (14), the element **-əŋ** shares the static locative-allative sense of the other two locative case clitics in the language.

2.2.2 ORIGIN OF CASE MARKERS: Where we know for certain the origins of a case marker in the Bodic languages, the source is invariably a noun. The historical succession of stages is roughly as follows:

- (15) *stage 1*: [N_a-GEN] N_b-LOC_c genitival modification of a locative-marked N
stage 2: [N_a] P_b(-LOC_c) locative-marked postposition
stage 3: N_a-LOC_b(-LOC_c) (compound) locative case

In stage 1 we have a genitive modifier of a noun in construction with a local case marker. By stage 2 the genitive usually disappears, and often so does the marker of local case on the new denominal postposition. By stage 3 the postposition becomes a case clitic. I'll provide two examples illustrating this succession of stages.

In Proto-Bodic there was a noun with a form something like ***s-naŋ** 'interior, inside' which likely derived from ***s-na** with a suffixed relational etymon ***Vŋ**.⁷ In the contemporary languages, this is widely attested as a noun, but it is also widely attested as a postposition and a case clitic, with and without the historic locative. All the stages find attestation from the Tamangic languages:

- (16) *stage 1*: CHANTYAL
 tʰim-ye nʰaŋ-ri
 house-GEN interior-LOC
 'to/in the house's interior'
- | | |
|---------------------------------------|--------------------|
| <i>stage 2</i> : THAKALI (Georg 1996) | NAR-PHU |
| tim ⁴ naŋ ³ -ri | tʰim nʰaŋ |
| house inside-LOC | house inside |
| 'inside the house' | 'inside the house' |
- stage 3*: CHANTYAL
 tʰim-nʰari
 house-INESSIVE
 'inside the house'

As discussed in Noonan (2008a), ***s-naŋ** has evolved into a comitative in a number of languages, and in others an allative [Central Monpa] and an instrumental [Ladakhi].

A rather more interesting case is Proto-Bodic ***(g-)lam** 'road, way', which has evolved into an ablative independently in Tamangic and Kiranti. A number of Kiranti languages attest both the nominal reflex and the ablative, the latter occurring either alone as the case clitic **-lam** or in combination with the widely attested case etymon ***ka**, whose uses in the Kiranti languages center around the ergative-instrumental-ablative syncretism, but which also signals generic locatives and other relational notions elsewhere [Noonan 2008a]. In the Tamangic languages, ***(g-)lam** is universally attested in its nominal form, and is also attested in Chantyal, Dhankute Tamang, and Seke as an ablative, in these languages either alone or in combination with their ergative-instrumental-ablative ***sa**. The semantic development is as follows: 'road' together with an instrumental marker is in construction with a place name resulting in a con-

⁷ Matisoff 2003 reconstructs ***na** 'ear', which, with the relational affix, may underlie ***s-naŋ**. Starostin & Pejros (*n.d.*) reconstruct ***naŋ**/***nak** 'inside' for Proto-Sino-Tibetan.

struction meaning ‘with/by means of the road (of) X’. Path senses are primary, with ablative senses developing from these. The old ergative-instrumental-ablative marker may either be retained as in Thulung **-laŋka** [Kiranti] and Chantyal **-gəmsə** [Tamangic] or deleted as in Dumi **-lam** [Kiranti] or Dhankute Tamang **-gjam** [Tamangic].⁸

2.3 DERIVATIONAL CASE COMPOUNDING: The transition from stage 1 to stage 2 is typically accompanied by the loss of the genitive, as illustrated in (16) above, but this is not universally the case, and such retentions account for all examples of derivational case compounding in Bodic. The most common path to derivational case compounding, under the definition given here, involves a grammaticalization chain whereby de-nominal postpositions governing the genitive lose their status as independent words and fuse with the genitive. This results in a declension in which the genitive may occur independently and may also occur in combination with case markers derived from the old postpositions.

An example of derivational case compounding can be seen in Kathmandu Newari (Hargreaves 2003):

- (17) -ya GENITIVE
 -ya-tɔ DATIVE
 -ya-ke COMITATIVE

In this language, the dative and comitative derived from denominal postpositions in construction with the genitive, the case they once governed. The genitive element has now become optional, rendering this example more typical of the usual progression from NP head [via postposition] to case marker in Bodic, which inevitably seems to involve the loss of the historic genitive.

Derivational case compounding in Bodic never achieves the sort of systematicity found in languages like Avar (Blake 1994), which combines a set of five basic static locational markers with a set of three directionals to yield a system of twenty cases: since the static locationals can occur independently as well [they are the base forms], this system can be analyzed as an instance of derivational case combining.

(18)

	STATIC LOCATION	MOTION TOWARD	MOTION AWAY	PATH
‘on top of’	-da	-d-e	-da-ssa	-da-ssa-n
‘at’	-q	-q-e	-q-a	-q-a-n
‘under’	-ʃ'	-ʃ'-e	-ʃ'-a	-ʃ'-a-n
‘in, among’	-ʃ	-ʃ-e	-ʃ-a	-ʃ-a-n
‘in a hollow object’	∅	-e	-a	-ssa-n

The Avar system with its combination of forms with local meanings bears some resemblance to typical instances of case stacking, which will be described in the next section.

⁸ In both Thulung and Chantyal, the erstwhile ergative-instrumental-ablative suffix may be deleted as well, making these examples more like their Dumi and Dhankute Tamang counterparts. And in Chantyal, the erstwhile ergative-instrumental-ablative may also occur alone as an ablative, though infrequently so apart from some common expressions: see below §2.4.

Kabak (2006) proposes that postpositions following inflected nouns are less likely to cliticize than those following uninflected nouns. The succession from stage 1 to stage 2 proposed in (15) has the genitive dropping out prior to the cliticization of the postposition [stage 3], and thus is not in conflict with Kabak’s hypothesis. The situation described for Kathmandu Newari is, however, as would, in principle, any instance of derivational case compounding whose origins lie in the cliticization of a postposition, *e.g.* the Tokharian A example in (2) and the Avar system described above, assuming that the directionals in the latter derive from postpositions. In any case, the typical loss of the genitive in the transition from stage 1 to stage 2 may explain the relative ease with which postpositions cliticize in Bodic and, taken together with Kabak’s hypothesis, the comparative rarity of derivational case compounding in this group.⁹

2.4 CASE STACKING: The frequency of case stacking in the Bodic languages varies greatly from language to language, though most branches provide evidence for the phenomenon, if only in the reconstruction of contemporary case forms. Informally, we can distinguish two sorts: case stacking for the expression of complex trajectories, and case stacking for the expression of other relational notions.

The first sort – case stacking for the expression of complex trajectories – is widely attested and in some languages is reasonably common. This is the case for the Tamangic language Chantyal, and complex examples involving up to three case clitics¹⁰ occur in spontaneous discourse:

- (19) *dhufɯ-phyaraŋ-mar-gəmsə*
 tree-SUPER-CIRC-ABL
 ‘from around the top of the tree’¹¹

Case stacking in Chantyal is used to code complex trajectories that cannot be expressed by single case clitics.

For the expression of complex trajectories, the order of clitics is fairly consistent. The rightmost clitic expresses a trajectory and the leftmost one a location; where there are three, as in (19), the middle one may describe either a static location or a motion. As a practical matter, the clitics marking trajectory in Chantyal case stacking are just two: **-gəmsə**, signaling source dynamic [ablative] senses, and **-mar**, signaling motion, random or circular, in the vicinity of something.¹² Since Chantyal instantiates the common tendency to use the same forms for static location and for goal dynamic [allative] senses

⁹ In principle, it is possible for a system of derivational case compounding to arise from case stacking where in an arrangement of $-LOC_1-LOC_2$ the first, $-LOC_1$, is used with more than one other case and particular instances of $-LOC_2$ come to be restricted to so as not to occur independently but only in $-LOC_1-LOC_2$ constructions. The Avar system may have its origin in such an arrangement, but I know of no cases in Bodic where derivational case compounding has arisen in this way.

¹⁰ It should be emphasized that these forms are all clitics [and not words] since together they form a single stress unit with their head noun.

¹¹ All of the case clitics in (19) are complex: **-gəmsə** was discussed in §2.2.2 above; **-mar** contains locative **-ri**; **phyaraŋ** contains the locative element **-ŋ**.

¹² **-mar** also signals location in the vicinity of something, as in (19).

[see Noonan 2005 for extensive discussion and exemplification], case stacking is not usual where the combination would signal goal dynamic senses.

In Chantyal, case stacking seems to be a productive process in the sense that novel forms may be produced and understood given an appropriate stimulus, but most instances in free discourse are probably lexicalized to some degree, and this opens the door for certain kinds of irregularity. For example, the ablative/path clitic **-gəmsə** has a number of variants which are in more-or-less free variation when the clitic is used alone with nouns, including the full form and two reduced forms: one involving the initial syllable **-gəm** only, and the other involving the final syllable **-sə**, itself a reflex of the old ergative-instrumental-ablative syncretism. In certain combinations with other case clitics, certain allomorphs are preferred.

- (20) -nɦari-gəmsə 'from inside'
 -mar-gəmsə 'from around'
 -phiriŋ-sə 'under and through'
 -phyaraŋ-sə 'over across'

Of the set in (20), the last example provides an interesting contrast: where the meaning 'over across' is intended, the reduced allomorph is required; there the mean 'from over, from the top' is required, the usual form is **-phyaraŋ-gəmsə**, *i.e.* with the non-reduced form. This is evidence for the specialization of the allomorphs, lexicalization, or, more likely, both. What is interesting is that the path senses evident in the meaning of the last two examples in (20) derive historically from the full form: recall the discussion in §2.2.2 where it was noted that the full form **-gəmsə** has its origin in a construction signaling path.

Case stacking for the expression of senses not involving complex trajectories is much less common and never productive, though examples can be found easily enough. For example in Spiti, a member of the Tibetan Complex (Sharma 1992), the dative [which expresses primary objects] can be signaled by **-phi-la**, which consists of benefactive **-phi** and dative-locative **-la**. In other words, the dative sense could already be signaled by dative-locative **-la**, but a new, dedicated dative, **-phi-la**, was created by case stacking. This use of case stacking – to create dedicated forms where the old forms express functional syncretisms – is one source for renewal of case systems in these languages.

2.5 SIMPLE HEADLESS ADNOMINALS:

It was noted in §1.3 that the sort of case compounding in the expression of simple headless adnominals, while seldom discussed in the literature, seems possible in all the Bodic languages I've investigated. I provide below another example, parallel to the example in (7) above, this time from Thangmi [Mark Turin, *p.c.*]:

- (21) gai-go-ye naŋ-ko-kai cah-Ø-u-no
 1S-GEN-ERG 2S-GEN-DAT bite-AG.SU.SG-3PAT-3→3.PRET
 'Mine bit yours' [*e.g.* of dogs]

Essentially, we have a bracketing like [[N-GEN]Ø-CASE], where the phrase-final case clitic simply attaches to the next available host. Looked at another way, the genitive func-

tions here as a sort of nominalizer, permitting the adnominal NP to fill a noun slot requiring another case. In principle, this could even be a genitive slot, though if it were, haplology would, at least in Chantyal, eliminate the second genitive, except in very careful speech:

- (22) nə-ye (<nə-ye-ye) p̄hale-gəmsə ka kha-m
 1S-GEN 1S-GEN-GEN leg-ABL blood come-NPST
 'Mine's leg was bleeding' [= 'My dog's leg was bleeding']

Notice that the English possessive adjective *mine's* is a doubly marked genitive, a rare instance of the phenomenon in English.

While case compounding of any sort is not that frequently encountered crosslinguistically, the constructions illustrated here have many parallels in languages that do not employ case compounding. For example, in the Romance languages, proper names deriving from place names and composed with the preposition *de* can be treated as simple nouns grammatically: *de d'Artagnan* 'd'Artagnan's'.

2.6 COMPLEX ATTRIBUTIVE NOMINALS:

As noted, the Bodic languages are characterized by employing a single grammatical form which is used both as a nominalization and as an attributive (Noonan 1997, 2007). In some languages, this form, which is conventionally referred to as a nominalizer, can be suffixed to a variety of lexical classes producing attributives that can also be treated as nouns. A few examples of attributive use from Chantyal follow:

- (23) təyla-wa saka
 yesterday-NOM ancestor
 'yesterday's ancestors'
- (24) yəwta dyammər-ma citro-ma-ye h̄ə-sərə-wa phəlphul-ma-ye raksi
 one dogwood-PL barberry-PL-GEN that-manner-NOM fruit-PL-GEN raksi
 'raksi from some fruits like dogwood and barberry'
- (25) syālk̄h̄ark̄ə-ōra-wa m̄ənchi
 Syalkharka-CIRC-NOM person
 'person from around Syalkharka'
- (26) chə məyna-ri-wa nani
 six month-LOC-NOM baby
 'six-month old child'
- (27) k̄h̄wara-ru-wa k̄əru
 wheat-COM-NOM hooded.barley
 'the hooded barley among the wheat'

When used as nouns, these forms can occur with any case clitic appropriate for their grammatical roles, resulting in a kind of case compounding:

- (28) chə məyna-ri-wa-ra syālk̄h̄ark̄ə-ōra-wa-ye n̄h̄in k̄ha-i
 six month-LOC-NOM-DAT Syalkharka-CIRC-NOM-GEN fear come-PERF
 'The six-month old was afraid of the one from around Syalkharka'

Such forms have much in common with the simple headless adnominals discussed in the last section in that the nominalization process creates a new noun that can take ad-

ditional case marking. This process can be recursive, as we note from example (10), repeated below:

- (10) na-sə məŋgəle-ri-wa-ma-siŋ-wa-ra dekhə-i
1S-ERG Mangale-LOC-NOM-PL-COM-NOM-DAT show-PERF
'I showed it to the owners from Mangale'

3. THE PREVALENCE OF COMPLEX CASE MARKERS IN BODIC

In this section, I will provide two crude measures indicating the overall frequency of complex case markers of any sort, either those resulting from case compounding or those arising from denominal postpositions and case clitics as discussed in §2.2.2.

The first such measure is taken from Noonan (2008a) and reproduced as Appendix 2. This chart provides a set of selected case etymons in Bodic, noting their occurrence with various case roles/meanings in the literature, their overall frequency in those roles, and the number of times they occur 'uncompounded'. 'Compounded' in this context could refer to any sort of case compounding or to stage 2 or 3 developments in (15) in §2.2.2: that is, to denominal postpositions or case clitics. The figures are drawn from the survey of 76 languages reported on in Noonan 2008a.

Appendix 2 shows that compounded forms occur with some regularity in Bodic, even in the expression of 'basic cases'. Note that the expression of complex trajectories via case stacking of the sort illustrated from Chantyal in §2.4 were not included; only forms described in the literature under the labels provided¹³ were counted.

Appendix 2 also shows that the meanings of these clitics are quite changeable over a long enough time span: see Noonan 2008a for discussion.

The second crude measure looks at a slightly larger sample of 80 Bodic languages and counts polysyllabic [consisting of two or more syllables] and monosyllabic forms coding particular case functions.¹⁴ The assumption here is that polysyllabic forms are more likely to have a recent polymorphemic origin than monosyllabic forms – though as an absolute principle this is, of course, false and is even false with respect to certain forms counted in this way in the data. Nonetheless, since the etymologies of most case markers in these languages are not known at this time, this crude measure can at least point to case meanings that are likely to be targets for reinforcement via complex forms of any sort. The results are displayed in (29), wherein 'percentage' refers to the percentage of polysyllabic forms found for any given case function, and 'number' refers to the actual counts: polysyllabic/monosyllabic.

¹³ The labels were reinterpreted and standardized for the purposes of the study.

¹⁴ It should be noted that both the 80 language sample and the 76 language sample referred to above are convenience samples drawn from available published and unpublished sources. They are not, nor do they pretend to be, genetically or areally balanced. The numbers and percentages given here should be interpreted in this light, namely as crude measures only of the phenomena discussed here.

(29)	CASE FUNCTION	PERCENTAGE	NUMBER
	comparative	61%	17/11
	ablative	34%	39/75
	comitative	32%	29/62
	dative	17%	16/77
	instrumental	6%	5/84
	genitive	3%	3/91
	locative	2%	3/125
	ergative	2%	2/86

[The numbers may be higher than the number of languages surveyed (80) because some languages have more than one form with a given function; not all grammatical descriptions provided a comparative, hence the low number count.]

The differences among the case functions are striking. For comparatives polysyllabic forms are the majority, and for ablative, comitative, and dative senses the percentages of polysyllabic forms are also relatively high. Instrumentals, genitives, locatives, and ergatives, on the other hand, are overwhelmingly monosyllabic. One way to interpret these data is to assume that those case senses likely to be coded by polysyllabic forms are the locus of new forms entering the case system via case compounding or via post-positions grammaticalizing to cases. Where the forms are overwhelmingly coded by monosyllabic forms, we can assume that these case senses are the ends of grammaticalization chains. A discussion of why the particular case functions array the way they do lies beyond the scope of this paper.

The data from these two crude measures both suggest that there are a substantial number of historically complex case markers in the Bodic languages and that the complex forms are unevenly distributed among the case functions.

4. FINAL REMARKS

This paper began with a typology of case compounding in the languages of the world and followed this with a discussion of the sorts of case compounding found in the Bodic languages. The final section discussed the prevalence of case compounding in Bodic, and the role of case compounding in the evolution of case-marking systems in these languages. One conclusion to be drawn from all this is that not only is case compounding frequently attested in Bodic, but it is a major factor in the historical development of case morphology, a fact that will have to be taken into account in any attempt to reconstruct the histories of these languages.

Abbreviations

1S	first person singular	INES	inessive
2S	second person singular	LOC	locative
3	third person	NOM	nominalizer
ABL	ablative	NPST	non-past
ACC	accusative	PAT	patient
ADES	adessive	PERF	perfective
AG	agent	PL	plural
ALL	allative	PRES	present
CIRC	circumlative	PRET	preterite
COM	comitative	SG	singular
COMP	comparative	SUB	subessive
DAT	dative	SUPER	superlative
ELAT	elative	→	direction of transitive relationship
GEN	genitive		
ERG	ergative		

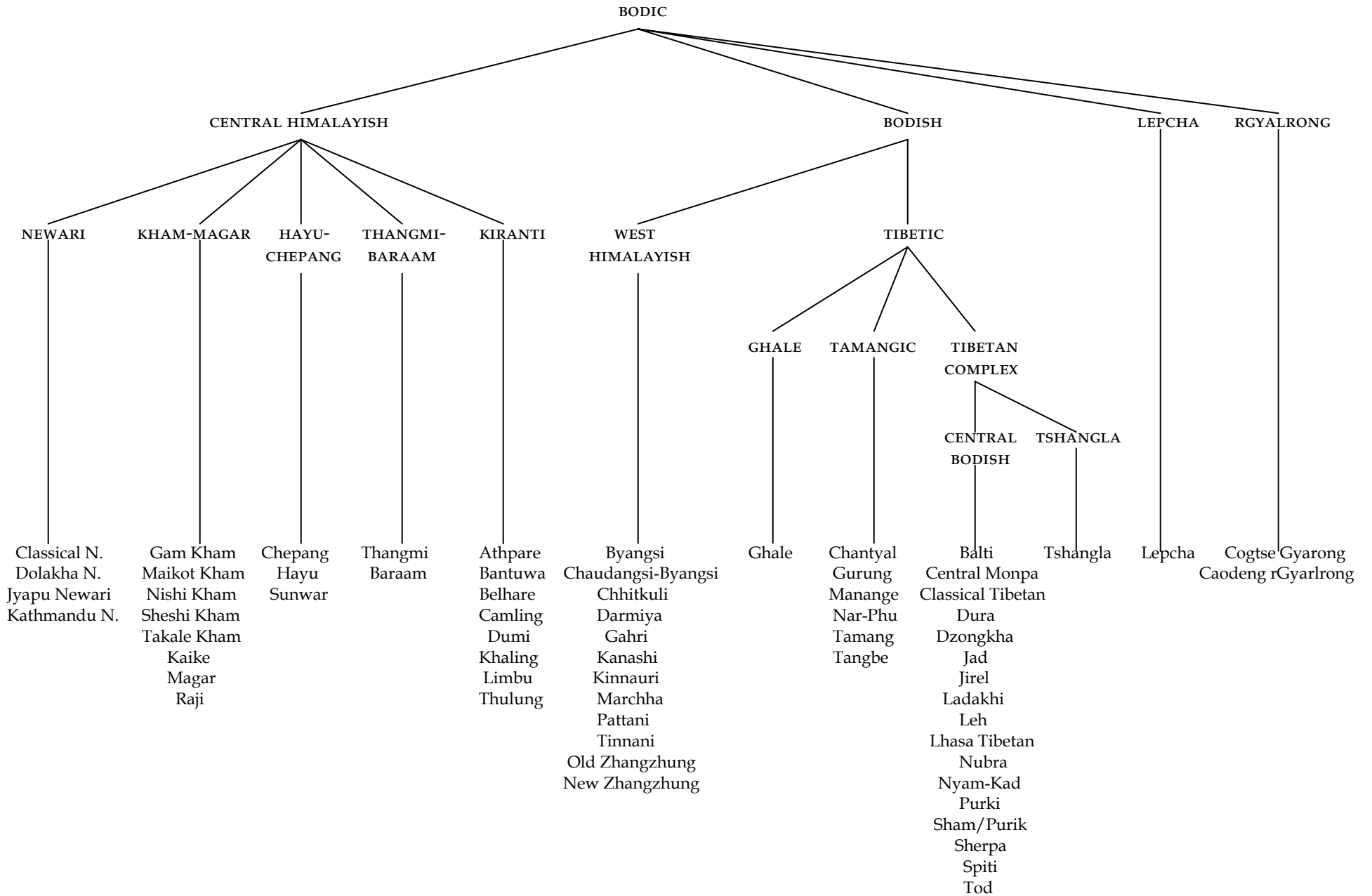
References

- Austin, Peter. 1995. 'Double case marking in Kanyara and Mantharta languages, Western Australia.' In Frans Plank, ed: *Double case: Agreement by Suffixaufnahme*, 363–379. New York: Oxford University Press.
- Bickel, Balthasar. 2003. 'Referential density in discourse and syntactic typology.' *Language* 79/4.
- Blake, Barry. 1994. *Case*. Cambridge: Cambridge University Press.
- Dryer, Matthew. 1986. 'Primary objects, secondary objects, and antidative.' *Language* 62/4:808-45.
- Evans, Nicholas. 1995. *A Grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Georg, Stefan. 1996. *Marphatan Thakali*. München/Newcastle: LINCOM EUROPA.
- Hale, Kenneth. 1982. 'Some essential features of Warlpiri verbal clauses.' *Work Papers of the Summer Institute of Linguistics, AAB Series A* 6:217–315.
- Hargreaves, David. 2003. 'Kathmandu Newar', in *The Sino-Tibetan Languages*, Randy LaPolla and Graham Thurgood, eds. London: Routledge.
- Huber, Brigitte. 2005. *The Tibetan Dialect of Lende (Kyirong): a grammatical description with historical annotations*. Bonn: VGH Wissenschaftsverlag.
- Kabak, Barış. 2006. 'An obstacle to the morphologization of postpositions.' *Studies in Language* 30.1:33-68.
- Koshal, Sanyukta. 1979. *Ladakhi Grammar*. Delhi: Motilal Barnasidass.
- Matisoff, James. 2003. *Handbook of Proto-Tibeto-Burman*. Berkeley: University of California Press.
- Moravcsik, Edith. 1995. Summing up Suffixaufnahme. In Frans Plank, ed: *Double case: Agreement by Suffixaufnahme*, 451–484. New York: Oxford University Press.

- Noonan, Michael. 1996. 'The fall and rise and fall of the Chantyal language,' *Southwest Journal of Linguistics* 15:1-2:121-36.
- Noonan, Michael. 1997. 'Versatile Nominalizations', in Joan Bybee, John Haiman & Sandra Thompson, eds. *Essays on Language Function and Language Type. In Honor of T. Givón*. Amsterdam/Philadelphia: John Benjamins.
- Noonan, Michael. 2003a. 'Chantyal', in *The Sino-Tibetan Languages*, Randy LaPolla and Graham Thurgood, eds. London: Routledge.
- Noonan, Michael. 2003b. 'Nar-Phu', in *The Sino-Tibetan Languages*, Randy LaPolla and Graham Thurgood, eds. London: Routledge.
- Noonan, Michael. 2005. 'Spatial reference in Chantyal,' in Yogendra Yadava, ed. *Contemporary Issues in Nepalese Linguistics*. Kathmandu: Linguistic Society of Nepal.
- Noonan, Michael. 2007. 'Nominalizations in Bodic languages.' In Elena Seoane & Maria Jose Lopez-Couso, eds. XXX. Amsterdam/Philadelphia: John Benjamins.
- Noonan, Michael. 2008a. 'Patterns of development, patterns of syncretism of relational morphology in the Bodic languages', in Jóhanna Barðdal & Shobhana Chelliah, eds. *The Role of Semantics and Pragmatics in the Development of Case*. Amsterdam: John Benjamins.
- Noonan, Michael. 2008b. 'Contact-induced change in the Himalayas: the case of the Tamangic languages', in Peter Siemund and Noemi Kintana, eds. *Language Contact and Contact Languages*. Amsterdam: John Benjamins.
- Plank, Frans, ed. 1995a. *Double case: Agreement by suffixaufnahme*. New York: Oxford University Press.
- Plank, Frans. 1995b. '(Re-)introducing Suffixaufnahme.' In Frans Plank, ed: *Double case: Agreement by Suffixaufnahme*, 3-110. New York: Oxford University Press.
- Schweiger, Fritz. 1995. 'Suffixaufnahme and related case marking patterns in Australian languages.' In Frans Plank, ed: *Double case: Agreement by Suffixaufnahme*, 339-62. New York: Oxford University Press.
- Sharma, D.D. 1992. *Tribal Languages of Himachal Pradesh. Part II*. Delhi: Mittal Publications.
- Starostin, Sergei & Ilia Pejros. *n.d.* [Reconstruction of Proto-Sino-Tibetan].
<http://starling.rinet.ru/babel.htm>

Appendix 1

Proposed Genetic Relationships Within the Bodic Section of Tibeto-Burman



Appendix 2

Distribution of the Reflexes of Selected Etymons

Numbers refers to the number of languages having a reflex of a given etymon with a given relational function.

	Erg	Inst	Abl	Gen	Dat	Loc	All	Com	Comp	Circ	Ines	Ades	Elat	Sub	Super	Path
<i>*ka</i>																
<i>overall usage</i>	1	3	11	12	4	19	3	2	1	1	2	1	1	2	1	5
<i>uncompounded</i>	1	1	1	8	3	14	2			1		1				1
<i>*ki</i>																
<i>overall usage</i>	9	7	3	14	1						1		2		1	
<i>uncompounded</i>	7	6	1	13	1										1	
<i>*(g-)lam</i>																
<i>overall usage</i>			7													7
<i>uncompounded</i>			3													3
<i>*na</i>																
<i>overall usage</i>	8	11	19	4	2	6	2	7	2		1					1
<i>uncompounded</i>	7	9	10	4	1	5	1	4			1					
<i>*nan</i>																
<i>overall usage</i>		1				1	1	6			23		2	1		
<i>uncompounded</i>		1				1	1	3			13			1		
<i>*nyampo</i>																
<i>overall usage</i>		2						10								
<i>uncompounded</i>								4								
<i>*Vŋ</i>																
<i>overall usage</i>		1	9	1	5	4		6	3		1	5		3	6	1
<i>uncompounded</i>			3		2	4					1	5				
<i>*r/la</i>																
<i>overall usage</i>		2	6	3	24	19	14	9	3	1	10	1	1	1	3	
<i>uncompounded</i>		1	3	3	21	14	13	1			6				1	
<i>*ri</i>																
<i>overall usage</i>			1		3	8	6				3	1	1	1		
<i>uncompounded</i>					3	8	6									
<i>*sV</i>																
<i>overall usage</i>	33	28	22	2		3	4	11	3		1		1			2
<i>uncompounded</i>	28	25	15	2		2	2	5								1