To appear in: David Bradley, Randy J. LaPolla & Graham Thurgood [eds.] *Language variation: papers on variation and change in the Sinosphere and Indospehere in honour of James A. Matisoff.* Canberra: Pacific Lingusitics

PROSODIC TAUTOMORPHEMICITY IN SINO-TIBETAN

BALTHASAR BICKEL

1. INTRODUCTION[1]

Sino-Tibetan is a prime example of how strongly a language family can typologically diversify under the pressure of areal spread features (Matisoff 1991, 1999). One of the manifestation of this is the average length of prosodic words. In Southeast Asia, prosodic words tend to average on one or one-and-a-half syllables. In the Himalayas, by contrast, it is not uncommon to encounter prosodic words containing five to ten syllables. The following pair of examples illustrates this.[2]

(1) Hakha Lai (Kuki-Chin; W. Burma)

($_{\omega}$ na-tuk) ($_{\omega}$ n̥aa) ($_{\omega}$ làay).
2SG.A-hit.with.stick 3PL.O FUTURE

'You will hit them.'

(2) Belhare (Kiranti; E. Nepal)

($_{\omega}$ mi-ŋŋ-u-ukg-att-u-n-chi-nn-^{h}ak=c^{h}a).
3NSG.A-NEG-roast-bring.down-PAST-3O-NEG-NSG.O-NEG-N=ADD

'They didn't even roast it for them down here.'

In the Southeast Asian language Lai, all but a few monomoraic prefixes license their own prosodic word (annotated here by 'ω'). The result of this in (1) is a sesquisyllabic word *natuk* 'you hit' followed by two prosodically and phonologically autonomous

---

[1] The research presented here was made possible by Grant No. 8210-053455 from the Swiss National Science Foundation. I am greatly indebted to Martha Paa (a.k.a. Kenneth Van Bik) and Bimala Pa (a.k.a. Lekhbahādur Rāī) for their help with Lai and Belhare, respectively. Parts of this paper were presented at the 29th International Conference on Sino-Tibetan Languages and Linguistics, Leiden, October 1996, and the 23rd Annual Meeting of the German Society for Linguistics, Leipzig, March 2001.

[2] Abbreviations are: A 'transitive subject', ADD 'additive focus', EXCL 'exclusive (of addressee)', O 'object', N 'nominalizer/focalizer', NEG 'negative', NSG 'nonsingular', PL 'plural', Q 'interrogative', SG 'singular'.

monosyllabic words *ŋ̥aa* ‘them’ and *làay* ‘FUTURE’. In example (2) from the Himalayan language Belhare, by contrast, no less than eight syllables are strung together into one polysynthetic word. There is but one main stress (on the first stem-containing syllable, here *ŋu*), and most morphemes undergo phonological adjustment when concatenated (Bickel 1996, 2002).

The monosyllabicity of Southeast Asian languages does not hold of GRAMMATICAL or LEXICAL WORDS, but of PROSODIC WORDS. Southeast Asian languages abound in bisyllabic or sometimes even multisyllabic compounds (often lexically frozen), and there is solid evidence that a sequence of morphemes such as the one in the Lai example (1), is just one grammatical word: the ordering of the object agreement marker *nhaa* and the future morpheme *làay* is absolutely rigid, their occurrence is strictly tied to the simultaneous occurrence of a verb, and no phrasal constituent can ever intervene between morphemes. These are all properties of bound morphemes, and this contrasts systematically with the behavior of autonomous grammatical words: Lai word order is relatively unconstrained and clause constituents allow re-ordering, interruption and independent use. Note that the same is true of words such as *will* and *them* in the English translation, although they are very similar in function to their Lai equivalents: *will* and *them* can be re-ordered (cf. *will you hit them?*), interrupted (cf. *you will definitely hit all of them*), and used independently (cf. *will you?*).

However, in languages such as Lai, grammatical and prosodic structure are not independent of each other: while one grammatical word may contain several prosodic constituents (feet or syllables), there is strong bias against the opposite, i.e., prosodic constituents (feet or syllables) containing more than one morpheme. I call this the TAUTOMORPHEMICITY PRINCIPLE, and the remainder of this paper is devoted to demonstrating that the effects of this principle are not only transparent in Southeast Asian Sino-Tibetan languages, but that they are also detectable in the polysynthetic and polysyllabic words of Sino-Tibetan languages in the Himalayas. There is evidence, then, that tautomorphemicity is not only an areally robust feature of Southeast Asia, but that at the same time, it is a genetically robust feature of Sino-Tibetan.

2. PROSODIC TAUTOMORPHEMICITY

The Sino-Tibetan Tautomorphemicity Principle is defined and further specified as follows:

(3) *The Tautomorphemicity Principle:*
Avoid feet or syllables containing concatenative morpheme boundaries.

The relevant prosodic constituents range over feet ('ϕ') and syllables ('σ'). The choice between these varies from language to language, but the claim here is that at least one of these two constituents disfavors internal morpheme boundaries in most Sino-Tibetan languages. The relevant boundaries are specified in (3) as concatenative (or 'agglutinative'). The principle does not apply to abstract morphemes that one might posit in accounting for tone patterns or stem alternations. Such patterns and alternations violate tautomorphemicity both necessarily and uninterestingly. Furthermore, the principle as set out in (3) is meant to capture a tendency, not a law. The precise degree and nature of its impact in any one language depends on a number of prosodic and morphophonological factors. Discussion of these is beyond the scope of this paper, but they can probably best be captured by Optimality Theoretic modeling (and see Bickel 1998 for a preliminary OT analysis of Belhare tautomorphemicity). In the following, I will put forward evidence for tautomorphemicity first on the syllable, then on the foot level.

2.1 Tautomorphemic syllables

The most obvious piece of evidence for tautomorphemicity on the syllable level comes from the lexical fact that most morphemes in Sino-Tibetan languages are at least one syllable long. This is generally true in Southeast Asia (Matisoff 1999), but it also holds for Belhare and other polysyllabic Sino-Tibetan languages. Lexical morphemes in Belhare all minimally contain one syllable, and the same is true for over 90% of the grammatical morphemes.[3]

One systematic exception to this is found scattered all over Tibeto-Burman and consists of the well-reconstructed coronal augments **s-, *-t and *-s*. Interestingly, however, more often than not these subsyllabic augments have lost their concatenative character and have instead developed into processes of phonation, tonation, or aspiration. Prefix **s-* is well-known to have often developed into aspiration (e.g., Lai *pit* 'be blocked' ~ *p^hit* 'block', Belhare *pok-* 'rise' ~ *p^hok-* 'raise'), and the suffixes **-t*

---

[3] Among about 80 such morphemes, there are only 3 clearcut exceptions: *-ŋ* '1SG.A', *-m* '1/2PL.A', and *-t* ~ *-ʔ* 'NONPAST'. Another exception is limited to one allomorph of a morpheme, viz. the postvocalic allomorph *-n* of the negation suffix *-ni*.

and *-*s* have often developed into phonation (e.g. Lai *fiaŋ* 'be clear' ~ *fiaʔn* 'make clear'; Peterson 1998).[4] A similar development is found with the Written Tibetan ergative -*s* which has developed into an umlaut and phonation alternation (e.g., WT *kho-s* 's/he-ERGATIVE' = /k^{h}ø̄^{ʔ}/, cf. absolutive /k^{h}ō/; Tournadre 1996). Through such developments, morphemes lose their concatenativity and become hidden inside the stem. There, they are shielded off against the effects of the tautomorphemicity principle.

Another systematic threat to tautomorphemicity arises from concatenating CVC with V-initial syllables. Simple syllabification of this would result in strings like ($_{\sigma}$ CV) ($_{\sigma}$ C-V), where the second syllable violates the tautomorphemicity principle. There seem to be two responses to this: onset-free syllabification and onset prothesis.

Onset-free syllabification can be exemplified by Garo (NE India), where the habitual tense suffix -*a* is always parsed as syllable-initial. Thus, *caʔ-a* 'eat-HABITUAL' is syllabified as ($_{\sigma}$caʔ)-($_{\sigma}$a), not as *($_{\sigma}$ca)($_{\sigma}$ʔ-a). Likewise, *kat-a* 'go-HABITUAL' is syllabified as ($_{\sigma}$kat)-($_{\sigma}$a), not as *($_{\sigma}$ka)($_{\sigma}$t-a). Structural evidence for this syllabification is (i) that there are no glottal syllable onsets in Garo, and (ii) that the /t/ in ($_{\sigma}$kat)-($_{\sigma}$a) does not undergo aspiration, a process limited to syllable initials in Garo (Burling 1961:5).

Onset-free syllabification is also found in the form of bisyllabic parsing of underlying vowel sequences. This is well attested in Himalayan languages. For example, in Dolakha Newar, a string like *ye-e* 'come-N' is realized bisyllabically, rather than as one syllable with a long vowel (Genetti 1994:30). Another instance is Belhare, where a sequence like *yu-a* 'go.down-IMPERATIVE' is realized by two distinct syllables rather than by one (raising) diphthong. The same is exemplified by the sequence *u-uk* 'roast-bring.down' in (2) above, which is syllabified into two nuclear peaks (on initial *ŋ*, see below).

The other response to CVC-V strings is onset prothesis: an underspecified onset (notated in the following by outline font) is inserted at the morpheme boundary. This segment is then realized together with the preceding consonant as an ambisyllabic geminate. Thus, CVC-V turns into ($_{\sigma}$ CVC)-($_{\sigma}$ ℂV), with an ambisyllabic Cℂ geminate. The result of this is that the underlying morpheme boundary comes to lie between rather than within syllables, as required by the Tautomorphemicity Principle. This process can be exemplified by Meithei (also known as Manipuri; NE India), where onset prothesis turns forms like *t^{h}əm-u* 'keep-IMPERATIVE' into ($_{\sigma}$t^{h}əm)-($_{\sigma}$𝕞u), or *təw-e* 'do-ASSERTIVE' into ($_{\sigma}$ təw)-($_{\sigma}$𝕨e) (Chelliah 1997:23, 67). Onset prothesis is

[4] On *-*t* in Belhare and other Kiranti languages, see below.

also sometimes recruited as a solution for hiatus (i.e., V-V) concatenations, which, too, are a potential threat to tautomorphemicity. In this case, the prothetic onset is typically a glide. An example of this are Belhare forms like ($_\sigma$ so)-($_\sigma$ yu) 'wait-3SG.O' (i.e., 'wait for him/her!'), rather than *($_\sigma$ sou̯), or ($_\sigma$ tu)-($_\sigma$ yu) 'dig-3SG.O' (i.e., 'dig it!'), rather than *($_\sigma$ tuu̯) or *($_\sigma$ tuː). In Belhare, the environments triggering glide insertion are lexically restricted. The default response to hiatus in this language is onset-free syllabification, as observed above with examples like ($_\sigma$ yu)-($_\sigma$ a) 'go down!'.

2.2 TAUTOMORPHEMIC FEET

In many Sino-Tibetan languages, tautomorphemicity holds of the basic stress unit, the foot (either in addition to or in stead of tautomorphemicity of syllables). In Southeast Asia, foot tautomorphemicity is often trivially satisfied by having a one-to-one correspondence between syllables, feet, and prosodic words. In languages such as Lahu, each morpheme is realized by a syllable that is a self-contained and equally stressed foot and indeed, word (Matisoff 1999). Foot tautomorphemicity is much less obvious in the Himalayas, but, as I argue in the following, there is one perspicuous morphophonological alternation that is best explained as an effect of just this principle.

We observed above that syllabic tautomorphemicity is sometimes ensured by means of onset prothesis. In Belhare this is found, as we saw, with CV-V sequences. Apart from this environment, onset prothesis also occurs in Belhare with CVC-VC sequences. This is illustrated by expressions such as *lap-uk* 'catch-bring.down', which is realized with a prothetic voiced consonant as ($_\sigma$ lap)-($_\sigma$ buk). Unlike unrepaired *($_\sigma$ la)($_\sigma$ b-uk), the form with onset prothesis complies with the tautomorphemicity principle.[5] An example from the Maivā-Mevā dialect of Limbu (Michailovsky 1986) is *huk-ɛn* 'hand-ARTICLE', which is realized as ($_\sigma$ huk)-($_\sigma$ kɛn) 'the hand', rather than as *($_\sigma$ hu)($_\sigma$ k-en). Again, onset prothesis makes the syllables tautomorphemic.

Onset prothesis is not found, however, in CVC-V sequences, although they are as much a threat to syllabic tautomorphemicity as CV-V and CVC-VC sequences; for instance, Belhare *lap-u* 'catch-3O' (i.e., 'catch it!') is realized as ($_\sigma$ la)($_\sigma$ b-u), not as *($_\sigma$ lap)-($_\sigma$ bu). In Belhare, only -VC morphemes trigger prothesis; heteromorphemic

[5] I have no explanation for the voicing, except perhaps that impressionalistically it yields a more pronounced demarcation of the second syllable, and as we will shortly see, stress unit. Place and manner features spread from the left.

syllables resulting from suffixing -V remain unrepaired. Unlike Belhare, Limbu has a phonemic vowel length opposition, and in this language, long-vowelled suffixes trigger prothesis as well: *kɛ-sira-t*^{h}*aŋ-i:* '2-pleasure-come.up-Q' is realized as ($_\sigma$ kɛ)-($_\sigma$ si)($_\sigma$ ra)-($_\sigma$ d^{h}aŋ)-($_\sigma$ ŋi:) 'do you like it?' (Phedāppe dialect; van Driem 1987: 144). This contrasts with short-vowelled suffixes: *kɛ-ni-siŋ-i* '2-see-REFLEXIVE-2PL' is realized as ($_\sigma$ kɛ)-($_\sigma$ ni)-($_\sigma$ si)-($_\sigma$ ŋ-i) 'you saw each other' (Phedāppe dialect; van Driem 1987: 384).

Thus, onset prothesis is generally restricted to -VC or -V: morphemes.[6] These morphemes carry both two moras, and this in turn is criterial for FEET in the Belhare and Limbu stress system. Stress in these languages follows a progressive trochaic pattern starting from the first stem mora. That the relevant feet are defined by bimoraicity, rather than bisyllabicity is evidenced by Belhare examples like *lab*^{h}*okarik* 'kind of small bird' which has secondary stress on ($_\phi$ rik): ($_\phi$ 'lab^{h}o)ka($_\phi$ ˌrik). If the stress system were syllable-counting rather than mora-counting, we would expect secondary stress to fall on *($_\phi$ karik): *($_\phi$ 'lab^{h}o)($_\phi$ ˌkarik). As the foot bracketing suggests, Belhare bans denegerate feet: the light syllable *ka* in this example remains unfooted and thereby unstressed (and indeed typically has a reduced realization). The only exception to the ban on degenerate feet is monomoraic syllables under word-initial (i.e., main) stress: ($_\phi$ 'mi) 'fire', or ($_\phi$ 'sa)(ˌmet). The Limbu stress system seems to be similar in these regards.

What this suggests is that the driving force behind onset prothesis in Limbu and Belhare is not tautomorphemicity of SYLLABLES but rather tautomorphemicity of the basic FOOT unit that underlies the stress system: ($_\sigma$ CVC)-($_\sigma$ ℂVC)- and ($_\sigma$ CVC)-($_\sigma$ ℂV:)- define sequences of two tautomorphemic bimoraic feet, i.e., ($_\phi$ CVC)-($_\phi$ ℂVC)- and ($_\phi$ CVC)-($_\phi$ ℂV:)-, respectively. By contrast, onset prothesis in CVC-V would create a monomoraic CV syllable unable to sustain a foot of its own. In brief, ONSET PROTHESIS OCCURS JUST IN CASE A TAUTOMORPHEMIC FOOT CAN BE OBTAINED.

As a consequence of this, onset prothesis is absent not only from CVC-V sequences but also from many other patterns. Consider, for example, a CVC-V-C sequence. Such a sequence has the same number of skeletal positions as the sequence that triggers onset prothesis (i.e., CVC-VC). Yet, in a CVC-V-C sequence, onset prothesis cannot possibly create tautomorphemic feet. Thus, strings like *lap-u-k=cha*

---

[6] There are three apparent exceptions in Maivā-Mevā Limbu: the negative imperative suffix *-ɛ*, the interrogative *-i* and the vocative *-e*. Although short-vowelled, these suffixes trigger prothesis. A likely explanation of this is that the behavior of these suffixes reflects an earlier VC or V: structure, as is evidenced by their Phedāppe cognates *-ɛʔ*, *-i:* and *-e:*, respectively (/e/ is redundantly long; there is no short /e/ in Phedāppe).

‘catch-3O-2A=ADD’ (i.e., ‘although you caught it’) remain unrepaired, for *($_{\phi}$ ˈlap)($_{\phi}$ ˌb-u-k)=cha violates foot tautomorphemicity as much as *($_{\phi}$ ˈla) ($_{\phi}$ ˌb-u-k)=cha. In the absence of segmental proliferation techniques other than onset prothesis, the only way to comply with foot tautomorphemicity in the case of CVC-V and CVC-V-C sequences is prosodic underparsing of syllables. Thus, forms like *lap-u* ‘catch-3O’ and *lap-u-k=cha* ‘catch-3O-2A=ADD’ are parsed prosodically as ($_{\phi}$ ˈla)b-u and ($_{\phi}$ ˈla)b-u-k=cha, respectively. ($_{\phi}$ ˈla) is a degenerate foot, but it occurs in word-initial position, i.e. the one position where Belhare tolerates degenerate feet (cf. above).

The result of underparsing is that there is no secondary stress in these examples. Suffix sequences show secondary stress only where foot structure is tautomorphemic. This is possible only with bimoraic suffixes, i.e. with -CVC (4a) or -VC (4b, repeated from 2 above) suffixes. Note that as long as the suffix is bimoraic it does not matter where in the suffix string the secondary stress comes to lie:

(4) Belhare

a. ($_{\phi}$ ˈta)r-u-c^{h}i-k-($_{\phi}$ ˌk^{h}ak)=c^{h}a
tar-u-c^{h}i-k-k^{h}ak=c^{h}a
bring-3O-NSG.O-2A-N=ADD
‘that you brought it to them as well’

b. mi-ŋ($_{\phi}$ ˈŋu)-($_{\phi}$ ˌuk)($_{\phi}$ ˌgat)t-u-n-chi-n($_{\phi}$ ˌn^{h}ak)=c^{h}a
mi-ŋ-u-uk-att-u-n-chi-n-^{h}ak=c^{h}a
3NSG.A-NEG-roast-bring.down-PAST-3O-NEG-NSG.O-NEG-N=ADD
‘They didn’t even roast it for them down here.’

Examples like these confirm the observation that (apart from the anti-hiatus cases mentioned in Section 2.1) onset prothesis in Belhare occurs in all and only those environments where the result is a tautomorphemic foot. This includes not only suffixal environments, but, as (4b) shows, also prefixal environments: in (4b), onset prothesis dissociates the stem-initial (and thus exceptionally degenerate) foot (ŋu) from the negative prefix *ŋ-*.

Table 1 summarizes the possible morpheme sequences of Belhare after CV and CVC stems. I leave out sequences with -CV and -CVC morphemes: -CVC morphemes license tautomorphemic feet to begin with; and -CV morphemes can easily be reconciled with the tautomorphemic foot principle by prosodic underparsing (non-stressing) of the kind noted just above. Note that in Belhare single C suffixes never follow a stem alone, without further affixes.

INSERT TABLE 1 ABOUT HERE (HORIZONTALLY)

The first two columns in Table 1 contrast the two patterns discussed earlier: CV-V sequences are brought in compliance with syllabic tautomorphemicity through either onset prothesis or onset-free syllabification, depending on what is lexically licensed by a given morpheme; by contrast, syllabic tautomorphemicity is not enforced with CVC-V sequences. But like CV-V, CVC-V is brought into agreement with FOOT tautomorphemicity. This is achieved through prosodic underparsing, i.e. non-stressing.

The second column illustrates V-initial bimoraic suffixes, which have the potential of projecting their own tautomorphemic foot. After CV morphemes, syllabic tautomorphemicity is achieved through onset-free syllabification (in the case of CV-VC). Foot tautomorphemicity is automatically ensured in this case. By contrast, CVC-VC sequences need onset prothesis in order to result in tautomorphemic feet. The resulting structure, CVC-ℂVC also satisfies syllabic tautomorphemicity but this is an epiphenomenon rather than the underlying reason because, as the first column of Table 1 shows, violations of syllabic tautomorphemicity alone, as in the case of CVC-V sequences, is no reason for onset prothesis.

The next two columns illustrate single C suffixes. Like before, the phonological result depends on what follows (and there is always a segment following single C suffixes in Belhare). If a single C suffix is followed by a monomoraic suffix, syllabic tautomorphemicity is left violated. Thus, like ($_\sigma$ CV)($_\sigma$ C-V) in the first column, both ($_\sigma$ CV)-($_\sigma$ C-V) and ($_\sigma$ CVC)-($_\sigma$ C-V) in the third column contain heteromorphemic syllables. Thus, just like before, violations of syllabic tautomorphemicity alone are not a sufficient reason for phonological repair. Examples of this involve the coronal augments *-t* and *-s* mentioned in Section 2.1: a string *hi-t-e* 'be.able-AUG-PAST' is realized as ($_\sigma$ hi)($_\sigma$ r-e) 's/he was able' (/t/ is regularly voiced and flapped between vowels). Foot tautomorphemicity is ensured, however, just like before, by underparsing: ($_\phi$($_\sigma$'hi))($_\sigma$r-e). The same pattern is found with -C-V sequences after CVC morphemes: *mat-t-u* 'narrate-T.AUGMENT-3O' is realized as ($_\phi$ ($_\sigma$mat))($_\sigma$t-u) 'tell the story!', with a heteromorphemic syllable but only a strictly tautomorphemic foot.

As for single C suffixes followed by -VC (last column in Table 1), the tautomorphemicity principle is complied with by means of C-deletion. The only environment for this again involves the coronal augment. (No other single C morpheme is ever followed by -VC.) In examples like the following, deleting the augment *-t* ensures tautomorphemicity of both syllables and feet:

(5) Belhare
n-($_\phi$ 'lu)-($_\phi$ ˌat)t-u-n
n-lu-t-att-u-n
NEG-tell-T.AUGMENT-PAST-3O-NEG
's/he didn't tell him/her.'

Incidentally, augment-deletion is also found before C-initial suffixes; the reason is again tautomorphemicity; cf. ($_\phi$ 'hi)-ma 'to be able to' from *hi-t-ma* 'be able-T.AUGMENT-INFINITIVE'.[7]

If the augment-VC sequence follows a CVC stem, as in *m-mat-t-att-u-n* 'NEG-narrate-T.AUGMENT-PAST-3O-NEG', C-deletion ensures tautomorphemicity of syllables, but at the same time it creates a sequence **mat-att*, which in turn corresponds precisely to the triggering environment of onset prothesis. And onset prothesis is indeed what one finds here:

(6) Belhare
m-('mat)-(ˌdat)t-u-n
m-mat-t-att-u-n
NEG-narrate-T.AUGMENT-PAST-3O-NEG
's/he didn't narrate it'

The result again confirms the observation that onset prothesis is found in all and only instances where a tautomorphemic foot can be created.

Other acccounts of onset prothesis are conceivable — e.g. by positing empty initial C-slots in -VC and -V: morphemes that are realized under certain lexical conditions; or by assuming that geminate-producing prothesis is triggered by morphemes that are pre-specified as ending in syllable codas (Michailovsky 1986) — but all these alternatives miss the generalization that the prothesis effect correlates with the distribution of stress and are therefore sensitive to foot structure.[8]

As noted above, in many Sino-Tibetan languages foot tautomorphemicity is satisfied trivially by having a one-to-one correspondence between morpheme,

[7] The classical analysis in Kiranti studies is that augments delete before C. This does not account for (5).
[8] For a more detailed critique of alternative accounts, see Bickel (1998).

syllable, foot and word; the principle does not require in this case any of the complex mechanisms found in polysynthetic languages like Belhare and Limbu. However, under one analysis, Southeast Asian languages recruit syllable underparsing as a means to comply with foot tautomorphemicity. Like in other Southeast Asian languages, the only deviation from strict monosyllabicism in Lai arises from prefixes like *na-,* as in *na-tuk* '2SG.A-hit.with.stick' in (1) above. Sequences like *natuk* are potential threats to foot tautomorphemicity since they can be analyzed in principle as unitary iambic feet. However, the fact that these prefixes are never stressed nor tone-bearing, and that they often have reduced vocalism (Matisoff 1999; cf. Chelliah 1997:38 on Meithei) suggests that these prefixes are not footed at all, i.e. that the foot structure is na-($_\phi$ tuk) rather than ($_\phi$ na-tuk). If this is correct, these languages would indeed recruit the same underparsing technique as found in Belhare and Limbu above.

3. CONCLUSIONS

The analyses presented in this chapter suggest that a seemingly disparate number of morphophonological alternations, stress rules, and historical developments receives a unified explanation once prosodic tautomorphemicity is acknowledged as an underlying principle of many Sino-Tibetan languages. The principle turns out to be not only an areally robust feature of Southeast Asia, but that at the same time, it is also found in a branch of Sino-Tibetan that is typologically and geographically far away from Southeast Asia: the Kiranti languages of the Himalayas. This suggests that tautomorphemicity is not only an areally but perhaps also a genetically robust feature of Sino-Tibetan: it not only spreads during language contact and language shift, but it also appears to survive language splits and family branching.

However, while prosodic tautomorphemicity shows strong areality in Southeast Asia, the principle seems to resist any convergence pressure on those Sino-Tibetan languages that are in intense contact with Indo-European in the Himalayas. Indo-European languages do not show any trace of tautomorphemicity; morphemes appear in any size, from single consonants to bisyllabic feet, and there does not seem to be any systematic technique for aligning syllable or feet boundaries with morpheme boundaries of the kind found in Limbu or Belhare. In turn, languages like Limbu and Belhare keep prosodic tautomorphemicity despite very intense languages contact in the form of large-scale bilingualism and code-switching with Nepali.

A reason for this diminished diffusibility in the Himalayas is perhaps that unlike in Southeast Asia, the effects of tautomorphemicity in Limbu and Belhare are very indirect: unlike in the case of the one-to-one correspondence of morphemes, syllables and feet in Southeast Asia, Himalayan tautomorphemicity manifests itself only indirectly through its effects on onset prothesis, phonation and stress rules. Unlike in Southeast Asia, there is no single and unitary tautomorphemic surface pattern as it were. It has sometimes been argued that the absence of such direct surface patterns makes linguistic features less transparent, and that this in turn may lower their areal diffusion potential (Heath 1978, Johanson 1992, Bickel 1999). Prosodic tautomorphemicity could be an instance of just such a feature.

If further research confirms this, the prosodic tautomorphemicity principle may prove to be a genetically highly stable feature of Sino-Tibetan — indeed, a genetic signature of the whole family. High diffusibility would then be expected only in the presence of monosyllabicism because this considerably increases the surface transparency of the principle.

References

Bickel, Balthasar, 1996, *Aspect, mood, and time in Belhare*. Zürich: ASAS Press.
1998, *Rhythm and feet in Belhare morphology*. Working Paper No. 287, Rutgers Optimality Archive; http://roa.rutgers.edu.
1999, *Grammatical relations, agreement, and genetic stability*. Ms., University of California, Berkeley; http://www.spw.unizh.ch/mitarbeiter/bickel/papers.
2002, Belhare. In Graham Thurgood and Randy J. LaPolla, eds *The Sino-Tibetan languages*. London: Curzon Press.
Burling, Robbins, 1961, *A Garo grammar*. Poona: Linguistic Society of India.
Chelliah, Shobhana L., 1997, *A grammar of Meithei*. Berlin: Mouton de Gruyter.
Driem, George van, 1987, *A grammar of Limbu*. Berlin: Mouton de Gruyter.
Genetti, Carol, 1994, *A descriptive and historical account of the Dolakha Newari dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
Heath, Jeffrey, 1978, *Linguistic diffusion in Arnhem Land*. Canberra: Australian Institute of Aboriginal Studies.
Johanson, Lars, 1992, *Strukturelle Faktoren in türkischen Sprachkontakten*. Stuttgart: Steiner.
Matisoff, James A., 1991, Sino-Tibetan linguistics: present state and future prospects. *Annual Review of Anthropology* 20: 469–504.
1999, *Genetic vs. areal linguistics in Southeast Asia: prosodic diffusibility in Southeast Asian languages*. Ms., University of California, Berkeley.

Michailovsky, Boyd, 1986, Structure syllabique et variation combinatoire: voisement et gémination en limbu. *Cahiers de linguistique asie-orientale* 15: 193–204.
Peterson, David A., 1998, The morphosyntax of transitivization in Lai (Haka Chin). *Linguistics of the Tibeto-Burman Area* 21: 87–153.
Tournadre, Nicolas, 1996, *L'Érgativité en tibétain*. Louvain: Peeters.

	-V(-C)	-VC	-C-V	-C-VC-
CV	σ: prothesis or onset-free syllabification ϕ: σ-underparsing $(_{\phi}(_{\sigma}\text{CV}))\text{-}(_{\sigma}\mathbb{C}\text{V})$ *or* $(_{\phi}(_{\sigma}\text{CV}))\text{-}(_{\sigma}\text{V})$	σ: onset-free syllabification ϕ: well-formed $(_{\phi}(_{\sigma}\text{CV}))\text{-}(_{\phi}(_{\sigma}\text{VC}))$	σ: unrepaired ϕ: σ-underparsing $(_{\phi}(_{\sigma}\text{CV}))\text{-}(_{\sigma}\text{C-V})$	σ: C-deletion ϕ: C-deletion $(_{\phi}(_{\sigma}\text{CV}))\text{-}\cancel{\text{C}}\text{-}(_{\phi}(_{\sigma}\text{VC}))$
CVC	σ: unrepaired ϕ: σ-underparsing $(_{\phi}(_{\sigma}\text{CV}))(_{\sigma}\text{C-V})$	σ: prothesis ϕ: prothesis $(_{\phi}(_{\sigma}\text{CVC}))\text{-}(_{\phi}(_{\sigma}\mathbb{C}\text{VC}))$	σ: unrepaired ϕ: σ-underparsing $(_{\phi}(_{\sigma}\text{CVC}))(_{\sigma}\text{C-V})$	σ: C-deletion ϕ: prothesis $(_{\phi}(_{\sigma}\text{CVC}))\text{-}\cancel{\text{C}}\text{-}(_{\phi}(_{\sigma}\mathbb{C}\text{VC}))$

Table 1: Overview of morpheme combinations in Belhare

Key:
first row in each cell: processes triggered by syllabic tautomorphemicity ('σ')
second row in each cell: processes triggered by foot tautomorphemicity ('ϕ')
third row in each cell: result