Areal and Universal Dimensions of Grammatization in Lahu

James A. Matisoff
Department of Linguistics
University of California, Berkeley

1. INTRODUCTION

The term grammaticalization, despite its heptasyllabic cacophony (the more concise grammatization, or even grammatization would be preferable) represents one of the most important phenomena in diachronic linguistics.

Grammat(ical)ization is inherently a diachronic concept. It refers to a historical semantic process whereby a “root-morpheme” with a full lexical meaning assumes a more abstract functorial or “grammatical” meaning. Such processes may take centuries to complete (as Bréal suggests: see below, but this does not seem necessarily to be the case. We can see grammaticalization phenomena occurring rapidly before our eyes in any language we choose.

Traditional Chinese grammar distinguishes between shí-chí “full/solid/real/true words” and xū-chí “empty words”, so that a good Chinese translation of grammaticalization is xū-huà (“empty-izing, emptification, evacuation”).

This metaphor has much to recommend it, since a hallmark of many grammaticized morphemes is that they are more conspicuous in absentia than they are in præsentia. Despite their “emptiness”, their absence can immediately stamp one as a non-native speaker, or at best, as a child. (Try saying, e.g. *I going soon, or *I read it in book.)

The pioneer semanticist (and reputedly the inventor of the word sémantique, Michel Bréal, conceived of grammatization in terms of specialization rather than evacuation:

We can see from this... example in what the Law of Specialization consists. Among all words of a certain kind, distinguished by a certain grammatical imprint, there is always one which is little by little drawn apart from its
fellow. It becomes the pre-eminent exponent of the grammatical conception of which it bears the stamp. But at the same time it loses its individual value, and is no more than a grammatical instrument, one of the wheels of the phrase. When the French say un temps plus long, une journée plus courte, the word plus serves to determine the adjective by which it is followed; but by itself it has no more existence from a semantic point of view than the inflection -ier. In the same way we can guess the reason why the Law of Specialization demands the help of centuries before it can operate. Words have too much individual significance to adapt themselves at once to playing the part of auxiliary. It is only a long use of words in various connections that slowly prepares the mind to remove from them their superfluous value. (Bréal 1897; trans. 1964, p. 15)

A more vivid characterization of the process is Givőn's term bleaching, which nicely captures the partial effacement of a morpheme's semantic features, the stripping away of some of its precise content so it can be used in an abstracter, grammatical-hardware-like way. The key notion here is abstractness (etymologically “drawing away”), a hard concept to define, since it seems to be almost a semantic primitive. One crude way of approaching it is in terms of what I have called juxtapositional productivity (Matisoff, 1969), an index of a morpheme’s combinatory possibilities in collocations — the fewer or more general the semantic features of a morpheme, the less likely they are to conflict with those of others. A full verb meaning ‘send someone on an errand’ will combine with fewer other lexemes than a bleached verb that means causative.

Grammatization may also be viewed as a subtype of metaphor (etymologically “carrying beyond”), our most general term for a meaning shift, or glissement sémantique. Grammatization is a metaphorical shift toward the abstract. Here G. Stern’s distinction (Bréal would say répartition) between the etymologically identical terms metaphor and transfer is useful: “metaphor” being defined as an originally conscious or voluntary shift in a word’s meaning because of some perceived similarity, while “transfer” is a meaning shift so natural that it happens unconsciously. Needless to say, the ordinary speaker is synchronically unaware of the semantic history of the functors of his language — such metaphors — (better, such transfers) are “dead”.

Human beings are preeminently creatures of habit, partly because of laziness, but partly because of the rational need to save effort. The scholar who eats the same lunch every day frees his mind from the time-consuming necessity of choosing something different each time. If a new collocation of morphemes proves useful (if it satisfies a Sprachnot), it can be stored in the brain as a unit, much like creating a “macro” key on a computer. Grammatized usages of a word may seem strange at first, but they can quickly begin to seem normal.

It could be claimed that all grammatization involves analogy — as indeed grammatical change in general seems to. Yet not all grammatical changes are instances of grammatization. One can think of many instances of ongoing grammatical change, where the younger generation accepts constructions which their elders find abhorrent — these are all analogical, though not necessarily grammatizational: e.g. bored of (instead of bored with), by analogy with tired of, sick of, so fun (instead of such fun), by analogy with, e.g. so amusing, so funny; between she and I (instead of between her and me), by analogy with, e.g. She and I are going); I could care less (instead of I couldn’t care less), perhaps by analogy with I should worry?; even the infamous new “double copula” construction (What I think is, is that he’s an idiot), perhaps by “analogical conflation” of a topicalization (What I think is — he’s an idiot) and an extrapolation (It’s that he’s an idiot).

It should also be pointed out that a morpheme may lose semantic features without becoming grammatized. The English expression white shirt was borrowed into Japanese as waishatsu in the Meiji period to mean “formal Western white shirt”, but was soon bleached (how appropriately!) to mean “Western shirt in general”, leading to such historically oxymoronic collocations as aoi waishatsu (lit. “blue white-shirt”). Similarly, the Lahu applied the expression Kàkà-phu (lit. “white Indian”) to the first “Europeans” they met, and this term has stuck; when they later encountered American blacks they unhesitatingly called them Kàkà-phu-nà (lit. “black white-Indian”). These are instances of generalization without grammatization. One could also call this idiomatization (see the discussion of Lahu “idiomatic verb concatenations”, 4.31 below).

What is really going on in such cases is a total loss of morphemic identity, or an obliteration of a morpheme-boundary. This goes far beyond what happens in grammatization, where the bleached item at least retains its morphemic autonomy. Examples of morphemic obliteration may be cited at will. When phonological reduction has not been too extreme we may talk of morphanization (Matisoff, 1973; 1978a), i.e. becoming an “orphan morph” (e.g. the cran- in cranberry or the -tril in nostril). More spectacular instances include the swallowing up of eye in window (< wind + eye) and daisy (< day’s + eye), and my all-time favorite, the -k in ink (ult. < Greek en-kaustos ‘burnt into’, with the second element < kai ein ‘burn’).
1.1. The Southeast Asian linguistic area

The Southeast Asian linguistic area, defined broadly to include China south of the Yangtze, NE India, and the Himalayan region, as well as peninsular and insular SE Asia, is one of the great linguistic areas of the world, with five major language families (Sino-Tibetan, Tai-Kadai, Hmong-Mien (Miao-Yao), Austroasiatic (Mon-Khmer), and Austronesian (Malayo-Polynesian) flourishing in symbiotic profusion. Despite their genetic diversity, long periods of intimate contact have largely homogenized all aspects of the phonology, grammar, and semantics of these languages, so that it makes eminent good sense to consider phenomena like grammatization in the widest possible areal context.

Although many SE Asian languages have more morphology than is generally realized, it is true that they are overwhelmingly analytic and (except for Austronesian) monosyllabic/tonal or sesquisyllabic (“syllable-and-a-half”) in structure. This is especially true of the languages that are genetically related to or culturally dominated by Chinese. In the latter category of “Sinospheric” languages are the unrelated Tai-Kadai and Hmong-Mien families, as well as Vietnamese (genetically Mon-Khmer but under intense influence from Chinese). The more than 250 Tibeto-Burman (TB) languages are demonstrably related to Chinese, but are different from it typologically, being (except for the Karenic branch) SOV — undoubtedly the original Sino-Tibetan word order.

Lahu, a language with half a million speakers on which I have been working since 1965, is typical of the Lolo-Burmese branch of Tibeto-Burman — the branch most firmly planted in the Sinosphere. (The TB languages of India and the Himalayas have more morphology, are less tonal, and are more under the cultural/historical influence of India than of China.) In this paper we will explore issues in Lahu grammatization that are of primary importance in the SE Asian linguistic area in general. The Sinospheric languages rely mostly on independent words rather than on affixes to express grammatical categories — particles, auxiliary verbs, classifiers, adverbs, etc. We will first consider the process whereby root-morphemes become grammatized into particles; then the dynamics of verb-bleaching in complex verbal nuclei; finally some phenomena which illustrate the truism that grammatical/semantic categories are continua rather than discrete, neat compartments.

All the Lahu examples are from my own fieldwork. The Thai and Mandarin examples are from standard sources, as is most of the Vietnamese, Khmer, and Hmong data (see References). The remainder of the latter, as well as the examples from Yao, are from my field-methods classes conducted at Berkeley.

2. “PARTICULARIZATION” IN LAHU

Particles are morphemes with abstract grammatical functions that cannot constitute the head of a phrase. Lahu, like other SE Asian languages, has dozens of them. They are conveniently divided into 3 subtypes: (a) Noun-particles (Pn) may only occur after nouns. Unlike true nouns, Pn’s may not be quantified or classified. (b) Verb-particles (Pv) only occur after verbs. Unlike true verbs, Pv’s may not be negated. (c) Unrestricted particles (Pu) may occur after either nouns or verbs. The Pu’s may be further subclassified according to whether they occur only in non-final clauses (non-final unrestricted particles or Pun’s), only in final clauses (final unrestricted particles or Puf’s), or in both final and non-final clauses (universal unrestricted particles or Puni’s). Note that in Lahu, as in the other SOV languages of the Tibeto-Burman family, all particles are postpositional.

Particles convey such essential grammatical information as case relationships, verbal aspect, inter-clausal propositional relations, evidentiality, the speaker’s emotional attitude, degrees of politeness, etc. Although as a group particles are obviously more abstract (or “empty”) than either nouns or verbs, within the particle class itself there is a hierarchy of abstractness. As one might expect, the “unrestricted” particles are the most abstract of all — their meanings being so general as to be congruent both with nominality and verbosity.

Sentence-final emotive particles are particularly richly developed in tone languages, since mere intonation is less salient when every syllable has a distinctive “melody” of its own.

Despite their centrality to the grammar of SE Asian languages, particles are generally open classes. Not only are particles readily borrowable from other languages, e.g. Lahu sê (Pv) ‘attitude of regret’ < Tai (cf. Siamese sê), they are also continually developing from fully lexical morphemes — i.e. particularization is an ongoing process. It is not always easy to say, in fact, where a root-morpheme ends and a particle begins. “In some cases the borderline between particles and members of other word classes is not too clear.” (Thompson, 1965: 258; below 4.4.)

Particles (along with other high-frequency abstract morphemes like adverbs and numerals) are often marked by special phonological features. In Lahu they are likely to lose their initial consonants: thâ? (Pn) ‘accusative’ is...
often pronounced ə (below 2.11); mā (Adv) ‘negative’ becomes simply ə; lè (Pun) ‘topicalizer’ varies with ə; tà or tā ‘perfective; durative’ (below 2.33, 4.42) frequently becomes ə or à, etc. Even if they manage to retain their initial consonant, particles may well lose their vowel quality through destressing: The Burmese negative adverb has been reduced to mə-, though it historically had the full vowel –ə (cf. Lahu mā); the Burmese numeral tə ’one’ (Written Burmese tə<PTB *g-tyik) is usually pronounced with a schwa (tə-). In extreme cases, a particle may get so reduced that it fuses with the preceding root-morpheme (cf. Lahu qay ’go’<*ka ’go’ + e (Pv) ’motion away from deictic center’; below 4.51). The limit of phonological slightness is reached when a particle comes to lack any segmental phoneme at all — it is convenient to treat an abrupt intonational contour on a Lahu verb as a verb-particle, “imperative glottal stop”, e.g. qay-p ‘Go!’ (Matisoff, 1973:353–4).

2.1. In the noun-phrase: A noun becomes a noun-particle

2.1.1. thə?, ‘upper surface’> OBJECT MARKER (direct or indirect); efficacy/agentivity depressant

This word has a good etymology that can be traced all the way back to Proto-Tibeto-Burman (PTB) *l-tək ʃ ’g-tək ‘ascend; above’ (see Matisoff, 1972a: #42; Benedict 1972: 52, 110, 123).

As an independent noun prefixed by ʒ-, or a constituent in noun-compounds, Lahu thə? means ‘upper surface; top part’:

ʒ-thə?–phə ʒ ’thə?-pə ‘the outer side; top surface’
khī–tə-qə-thə? ‘instep; upper side of foot’
lī? ʒ-thə? ‘on the book; the top surface of the book’

(1) ʃə-ʃə chi mī-chə ʒ-thə? bu te a sand DEM floor surface pile put PTCL
‘Pile the sand up on the floor!’

As a noun-particle (often reduced to hə? or even ə?) (see above), this word has developed into an object-marker (Matisoff, 1973: 155–8). In this function it is used rather sparingly, only where clarity demands or emphasis is required. When both direct and indirect objects are present, thə? will follow the indirect object. This is because indirect objects are typically human (i.e. potential agents, high on the scale of efficacy), so that an explicit non-agent marker is sometimes required to exclude such an interpretation:

(2) lī? chi šə thə? pī tā ve yō book DEM 1SG *** give PTCL PTCL PTCL
‘(Someone) has given me that book.’

Without thə?, (1) could be interpreted as “I have given (someone) that book.” This Pn may also be used in a temporal sense, marking a sort of “accusative of time”:

(3) ʒə-hi khə də? ve qə̃? cə ve ha-pə thə? yō 1PL top hit REL year eat GTN month *** PTCL
‘We play with tops in the month that we celebrate New Year’s.’

thə? may also mark embedded sentential objects:

(4) yə qə̃? la ve thə? ʒə də-lə ve yō 3SG return PTCL PTCL *** I hope PTCL PTCL
‘I hope that he comes back.’

A clue to the semantic development of this etymon is found in the archaic Red Lahu ritual language, where one finds a compound noun-particle ú-thə? with a general locative or accusative meaning. The first syllable here is undoubtedly to be identified with ú- ’-‘head’ (cf. ú-gə ’pillow’, ú-qo ‘head’):

(5) cho-ú-yə-ú-thə? ʒə mə qa people *** affect NEG reach
‘May (these misfortunes) not be visited upon the people!’
(cf. Eng. ’upon the heads of the people’)

(6) mú-ni cə na ú-thə?, bə-pə cə na ú-thə? sun be above *** moon be above ***
in the place above the sun, in the place above the moon’

It is easy to see how the basic meaning ‘top surface’ could have been bleached into the abstract idea of ‘impingement upon; being topped by the action of the verb.’

2.1.2. lo ‘road/way’ > locative particle

Unlike the previous example, where the fully lexical meaning of the etymon continues to coexist alongside its grammaticized one, the original meaning of the locative Pn lo can only be deduced from comparative data.

There is a well-attested PTB etymon *lam ’road; way’ (STC #87) (cf.
Written Tibetan (WT) lam, Written Burmese (WB) lăm, which has not survived with this meaning in Lahu. (The Lahu word for 'road' is now yâ'-qə, perhaps related to the verb yâ? descend'.) Instead, Labu has developed a locative Pn lo, which I believe to be a variant of the 'road' root. -o is the regular Lahu reflex of *-am (cf. WB lam, Lahu lâ? 'cord (measurement)'); the Lahu mid-tone, unmarked in my transcription, points to Proto-Lolo-Burmese (PLB) Tone *3, while the WB cognate reflects Tone *2, but this is no problem. Tonal developments are frequently irregular with functors, and Tone *3 is demonstrably secondary in any case.

As a Pn, lo is somewhat more formal than the other locative particles 5 and kâ?. By itself lo does not specify direction of motion, or even motion vs. rest; the interpretation depends on the following verb, or the sentence as a whole:

(7) há-qə lo mi chê ve
cave *** sit PROG PTCL to be 'He's sitting in the cave.' ("ESSIVE")

(8) há-qə lo là? e ô
towards *** enter PTCL PTCL 'He has already gone into the cave.' ("ADESSIVE")

(9) há-qə lo là? e ô
*** emerge from 'He has already come out of the cave.' ("ABESSIVE")

Convincing comparative evidence for this grammaticalization comes from an obscure, moribund language of E. Nepal, Hayu (Michailovsky 1989), which has a cognate noun lom 'way; road', as in lom he (N + Pn) 'sur le chemin; on the way' and ek din mó lom ("one + day + GEN + road") 'un jour de chemin; a day's journey'. In Hayu also this noun seems already to have developed into a noun-particle, as in (10):

(10) cə ang lom na jú!
ISG *** EMPH descend 'Come down my way!'

Sequences like ang lom have changed their internal modificational structure, passing from a MODIFIER + HEAD (where the pronoun is genitively attributed to the following noun: my way) to a HEAD + MODIFIER (where the bleached noun is the satellite of the preceding pronoun: first person locatively conceived).

2.1.3. 'object\possession' → genitive marker

Thai, Vietnamese, and Cambodian have all developed genitive markers out of a noun meaning "thing: object":

a. Thai khăơŋ
   As noun:
   (11a) paj sûy khăơŋ
         vV Vh ôbi
   'go buy things'
   As particle:
   (11b) mia khăơŋ phôm
         wife GEN 1SG 'my wife'

b. Vietnamese cía
   As noun:
   (12a) cía cóng
         Nh public
         'public funds; state property'
   As particle:
   (12b) cái va-li cía óng Nam
         CLF valise GEN Mr.
         'Mr. Nam's suitcase'

c. Khmer rəph
   As noun:
   (13a) rəph nhū kee hav thaa kmaw-day
         thing DEM 3SG call QUOT pencil
         'That thing is called a pencil.'
   As particle:
   (13b) pu-qúaq tòuc rəph kôm pìi neəq nih
         friend little GEN 1SG two CLF DEM
         'these two little friends of mine'

2.1.4. 'female proprietary spirit' > feminine agentive nominalizer

The original or basic meaning of Lahu (g-)lē-ma seems to have been 'female proprietary spirit', e.g. the guardian spirit of a stream. The syllable
še is probably from the same etymology as qhâʔ-še ‘village headman’. še (V) ‘lead’ (< PLB *sîh1 šè *sîh2 or *sin1 šè *sin2; cf. WB soû, *soû ‘owner: proprietor’); -ma is a feminine suffix.

From there the meaning branched out in several directions. In composition with the morpheme pê ‘house’ (pê-še-ma) it developed two separate meanings: (a) the more “literal” one, lady of the house, and (b) the more idiomatic one ‘house lizard’, a harmless, ubiquitous creature that lives on the ceilings and walls of every house in Thailand. They are called cîyêk in Thai, and, like our ladybugs, are regarded with some affection.¹ From (b) the meaning developed into ‘bug; critter’ in general (e.g. pí-še-ma ‘silkworm’), and very recently the missionaries have encouraged a neologistic development into ‘microbe, germ’ (probably a calque on English, e.g. a flu bug). All this is certainly not grammatization, but merely glissement sémantique, metaphor, idiomatization.

On the other hand, the notion ‘female proprietary spirit’ developed more mundanely into ‘female owner in general’, and (in the direction of inalienable possession) into ‘the female body’:

(14a) yê 3-še-ma dàp jà 3SG FEM.body good very
She’s got a great body.

From there it was only a small further step to two thoroughly grammatized usages:

(a) feminine reflexive

(14b) yê 3-še-ma kán chi qha pa-è ña te ve yô 3SG *** work this all must do PTCL PTCL
She has to do this work all by herself.²

(b) feminine agentive nominalizer

(15) yâ ña ga pa pì 3-še-ma ‘midwife’
(i.e. “the who must help give birth”³)

This is an interesting example of the same lexeme developing both highly concrete and highly abstract senses.

2.2. Verb > noun-particle: ‘Copula’ > pluralizer

Here is a case that can only be “teased out” as the result of much comparative/historical labor. In “God and the Sino-Tibetan copula” (1985), I traced the phonological and semantic evolution of a PST etymon *s-ray šè *s-way, which originally seems to have had a highly abstract copular meaning. The variant in -w- gave rise to the most important and complex Lahu particle we, which now does duty as a nominalizer (as such it occurs in the citation-form of verbs, like Eng. to) as well as a subordinator (relativizer, genitivizer) (see Matisoff, 1972b). The variant in *s-r-, on the other hand, developed into a pluralizer in various TB languages — maximizing either the subject or the object of the verb:

Lakhru krai ‘plural affix for denoting company, etc.’; Mikir haî ‘plural of non-animate nouns’; Boro sray ‘to V everything’ (e.g. za-sray ‘eat everything’, zankri-sray ‘shake everything’, ran-sray ‘distribute everything’).

In Lahu, the descendant of this variant of the etymology is the morpheme -hi, used as a pluralizer with the personal pronouns and some proper nouns or nouns referring to people: ñà-hi ‘we’, nà-hi ‘you (pl.)’, yô-hi ‘they’, Cà-bû-hi ‘Cà-bû and his friends; Cà-bû and his group’.⁴

2.3. In the verb-phrase: From verb to verb-particle

This sort of development is highly typical of SEA languages in general. The pan-areal criteria for distinguishing full verbs from verb-particles are negatability and occurrence as the only morpheme in a VP. The presumed development is from full verb to verb-particle. I have found no instances of the development proceeding in the opposite direction, i.e. “degrammatization” of particles to full verbs.

A neat example of the difference is the behavior of the Thai morpheme ?dàt, which behaves sometimes as a full modal verb meaning ‘be able to; capable of’ (in which case it may be directly preceded by the negative adverb mây), and sometimes as a modal particle meaning ‘likely to; imminent possibility’, in which role it may not be directly negated:

(16a) mây ?dàt cô pàj NEG Vmod PTCL Vh ‘isn’t able to go’

(16b) ?dàt cô mây pàj PTCL PTCL NEG V ‘probably won’t go’
(In the terminology of Noss 1964, a modal verb is one which can precede another V in the same verbal nucleus. A modal particle also precedes the verb it modifies.) Similarly, in Chura (a Mon-Khmer language of the Bahnaric group, spoken in Vietnam), where “auxiliary verbs are distinguished from preverbal particles in that the auxiliaries are negatable” (Thomas, 1971: 144–8), the word conh may function either as a full, negatable “intentional verb” meaning ‘want to’ (17), or as a non-negatable preverbal particle meaning ‘almost, about to’ (18):

(17a) anh conh saq
    1sg Vint Vh
    ‘I want to go’

(17b) anh éq conh saq
    NEG
    ‘I don’t want to go.’

(18) anh conh churt
    PTCL V
    ‘I’m about to die’

But not:

*(18a) anh éq conh churt/
    NEG PTCL V

The same semantic connection between imminence and volition is also illustrated by the range of meanings of the Lahu particle gá (see 2.31 below and note 5), and the Burmese particle hci, as well as by the Yao verb/verb-particle ñy — though we are here dealing with a meaning constellation that is more universal than areal, as witness colloquial English expressions like “Looks like it wants to rain. doesn’t it?”

2.3.1. ‘think/desire’ > desiderative particle

In modern Lahu, the important morpheme gá combines freely with verbs to form desiderative expressions:

qay gá ‘want to go’, ší gá ‘want to know’, gê gá ‘want to laugh’

Yet gá is not itself a verb, since it cannot be preceded by the negative adverb mà (mà qay gá ‘not want to go’, NOT *qay mà gá), nor can it constitute a verbal nucleus all by itself. To translate an English sentence where ‘want’ is the main verb (e.g. “I want mangoes”), a Lahu verb must be supplied for the gá to hitch onto:

(19a) ṭà má-mo-si hë? gá
    ISG mango get PTCL
    (lit. “I want to get mangoes.”)

(19b) ṭà má-mo-si cå gá
    eat PTCL
    (lit. “I want to eat mangoes.”)

Other occurrences of gá give clues as to what its original non-grammatized meaning must have been. Besides its use as a desiderative particle, it also occurs as a bound morpheme in “elaborate expressions” (an elaborate expression is a 4-syllable collocation formed with elegant intent, where either the 1st and 3rd, or the 2nd and 4th syllables are identical), where it appears as the “couplet” of the free verb dcá ‘think’. An elaborate couplet consists of the non-repeating syllables in an elaborate expression, usually synonymous, antonymous, or otherwise correlative in meaning, thus:

dcí-phu-gá-phu ‘pure thoughts’ (‘think-white-gá-white’)

dcí-yé-gá-yé ‘be strong-minded; have equanimity’
    (‘think-firm-gá-firm’)

dcí-sha-gá-sha ‘be serene; have peace of mind’
    (‘think-easy-gá-easy’)

Certain expressions in the archaic ritual language clearly indicate that gá must once have had an independent verbal meaning like “think of conatively; desire”:

gá-pi-gá-së “desire-bundle” (ritual object offered in the temple, consisting of a bundle of sticks tipped with cottonwool to simulate flowers).

2.3.2. ‘come’ > cisative particle and marker of non-3rd person beneficiary

Sometimes the grammatization of a Lahu root-morpheme is accompanied by a morphophonemic change — notably a change in tone (which may be the synchronic reflex of a now-vanished segmental affix). The full verb ‘come’ is in the low-falling tone, là (e.g. mà-yé là ve ‘It’s raining’ (“rain comes”)). This verb has been grammatized into two semantically related particles, each of which is under its own distinctive tone: la (mid-tone (unmarked in the transcription)) and là (high-falling tone). 6
The mid-toned particle la indicates cisativity, either in the "literal" sense of motion toward the deictic center, or figuratively, in the sense of "almost coming to pass; nearly V", e.g. mā? la (V + Pv) "blow in this direction; blow hither", pō la 'come flying', cā la 'come and eat', sī la 'be close to death', cā la 'come to possess; get a hold of'.

The falling-toned particle là has an even more abstract meaning. It is a benefactive particle indicating that the verbal action is for the benefit of (or impinges upon) a non-3rd person, e.g.:

cho là (V + Pv) 'chop for me/us/you'
phē là (V + Pv) 'release me/us/you'
phā?-mā là (V + Pv) 'confess to me/us/you'

In this usage, là forms a tight paradigmatic opposition with a grammatized verb pī 'give', which indicates that the verbal action impinges on a 3rd person, e.g.:

cho pī (Vh + Vv) 'chop for him/her/them'
phē pī (Vh + Vv) 'release him/her/them'
phā?-mā pī (Vh + Vv) 'confess to him/her/them'

See Figure 1.

The highly grammatized nature of the opposition between là and pī is noteworthy, despite their different grammatical statuses (là is a Pv, tonally differentiated from its parent verb, while the pī is homophonous with its parent verb, so must be regarded as a post-head versatile verb ("Vv"; below 3.1). (For a brief discussion of the causative function of pī, see below 3.2.3.)

2.3.3. 'place/put down' > perfective, durative, permanent action

The Lahu verb tā 'place on; put on; set on; stand sthg. upright' has impeccable Tibeto-Burman credentials, reflecting PLB *tiʔa? (cf. WB thā) and ultimately PTB *s-ia (cf. WT sta ʔ stad):

(20) pē-hā tā ve
candle ***
'Set up a candle.'

(21) ni-mā tē sī ti tā ve
heart one CLF only *** PTCL
'concentrate one's efforts; be united in spirit with others'

This etymon has been bleached into one of the most important Lahu Pv's, in which function it indicates that the verbal action is stable, quasi-permanent, likely to last (durative); or that it is over and done with (perfective):[10]

(22) yā-hi å-quo ti chè tā ve
3PL home only stay *** PTCL
'They just stayed at home.'

(23) cī tā ve vāʔ-å-qō
boil *** GEN pig-head
'a pig's head that has been boiled'

The particle tā does not collocate with adjectival verbs, apparently because of pernicious redundancy (adjectives already refer to quasi-permanent qualities or states). This fact provides one of the best criteria for distinguishing
the class of verbal adjectives from the other verbs in the language (adjectives are a subclass of verbs by virtue of their negatability (above 2.33)):

**ORDINARY VERB**

-mā qay 'not go'

**ADJECTIVAL VERB**

-mā dā? 'not good'

qay tā 'had gone; went for good'

*dā? tā* (nonsense)

This morpheme is straddled precariously on the interface between versatile verbhood and particlitude. (See below 4.2.1.1, 4.4)

2.4. Verb becomes a clause-particle (complementizer)

Often a verbal morpheme comes to govern a whole clause, in the manner of a subordinating conjunction. Of the many examples which could be given, we confine ourselves here to a couple of typical developments. (We will return to this topic under "Causatives", below 3.23.)

2.4.1. *say* > quotative/topicizer/conditional

2.4.1.1. Thai wāa 'say; quotative complementizer'

The Thai word wāa still retains its full verbal meaning 'say' in certain contexts:

(24) wāa phōm duu-thùuk nām, māj ciŋ loŋ

* 1SG despise DEM NEG true EMPH

'To say that I look down on them is simply not true.'

At the end of a non-final clause containing a verb of utterance or of cognition,11 before the intonation break, this morpheme appears in grammaticized form as the introducer of the quotative clause — i.e. the clause that specifies what has been said or thought about:

(25) phōm kā jay māj nēe-caj wāa, sa

1SG conj still NEG sure *** PTCL

paj dā? rāy māj

go able or not

'I'm still not sure whether I'll be able to go or not.'

Among the large number of verbs that may take a quotative wāa-complement are:

(a) **UTTERANCE:**

phūt 'speak; say', bōk 'tell', riák 'call', sānjuā 'promise', tamnī-ti?-ti? 'blame; criticize', etc.

(b) **COGNITION:**

riū/sāap 'know', khit 'think', rū-syék 'feel' nēe-caj 'be sure', dājjin, 'hear', wāy 'hope', etc.

Sometimes there is a slight difference of emphasis, according to whether the wāa is taken in its more verbal, postpausal sense (26a), or its more grammaticized, prepausal sense (26b) (see Noss 1964, p. 200):

(26a) khāw phūt samā, wāa ?aahān māj ?ərəj

3SG speak always *** food NEG tasty

'They're always talking (about it, saying among other things) that the food is no good.'

(26b) khāw phūt samā wāa, ?aahān māj ?ərəj

'They always say that the food is no good.'

2.4.1.2. Khmer thaa 'say; quotative complementizer'

As a main verb Khmer thaa means 'say' (27). As a quotative complementizer (28, 29) it behaves exactly like its Thai counterpart, wāa:

a. As main verb:

(27) look thaa mac

2SG *** how

'What did you say?'

b. As quotative complementizer:

(28) kāom kit thaa look qayuq prahael mophi? pram

1SG think *** 2SG age about twenty-five

'I think that you're about 25 years old.'

(29) rabok nih kee haw thaa qway

thing this 3SG call *** what

'What do they call this?' ("This thing they call saying what?")

(See (13a), above.) This Khmer expression haw thaa qway is exactly analogous to Thai riák wāa ?ərəj, also meaning literally "call-say-what".)
2.4.1.3. Lahu qoʔ ‘say’ > qo ‘if’, qoʔ qo (‘if one says’) ‘topicalizer.’

The Lahu verb qoʔ ‘say; tell; call by a name’ is routinely used to mark off a preceding clause as quoted material:

(30) yò kàp mà qay gà qoʔ ve
 3SG PTCL NEG go want *** PTCL
‘He says he doesn’t want to go either.’

The etymology of this verb is not certain, but it seems definitely to be related to a mid-toned particle qo ‘topicalizer; if, when’:

(31) phi chi tê gii qo a-θ6-ma te tu le
dog DEM one CLF *** what do PTCL Q
‘As for this bunch of dogs, what’ll we do (with them)?’

(32) nò òe-ve că qo, nà pà ve yö
2SG DEM eat *** sick PTCL PTCL PTCL
‘If you eat that, you’ll get sick.’

(33) ha-pa chi tê fi pa-ò qo
month DEM one CLF end ***
‘When this month is over’

Speakers seem to retain an instinctive feel for the relationship between qoʔ and qo, since they are sometimes used together to form a more emphatic topicalizer, lit. ‘if one speaks of CLAUSE’:

(34) lò hé ve qoʔ qo 5, a-yan ga
book study PTCL *** PTCL time must
pà mà ve yö
spend much PTCL PTCL
‘When it comes to studying now, you’ve got to spend a lot of time at it.’

2.4.1.4. Tamang 3pi-sam and Mandarin de huà

The development of topicalizers or conditionals from verbs meaning ‘say’ is documentable in many other languages of East and Southeast Asia: In Tamang (a TB language of Nepal), the sequence of morphemes 3pi + sam (lit. SAY + IF “if one says”) behaves as a unitary conditional marker, exactly like Lahu qoʔ qo:

(35) 1nì-ci 3pi-sam
   ‘s’il est parti’,

‘..où le verbe dire, 3pi, tend à jouer le rôle d’une particule, d’un monème grammatical’ (Mazaudon, 1973:50).

In Mandarin, the noun huà ‘speech’ may occur as the head of a relative clause (marked by the preceding subordinating particle de), with the whole clause then functioning as the protasis of a conditional sentence:

(36) tâ yé hù kân lái de huà,
3SG PTCL NEG willing come ***
wòmen zênme bún ne?
1PL how do PTCL
‘If he’s not willing to come either, what shall we do?’

(Literally the first clause means something like “(as for) the talk of his also not being willing to come”.)

Yet here too we seem to be dealing with a universal grammaticalization tendency, as witness the English use of say to introduce hypothetical clauses: “Say you give me two coconuts…”

2.4.2. ‘come’ > purpose marker

The Chinese verb lái ‘come’ plays many special grammatical roles in the language. As an auxiliary after verbs of motion, it functions like Lahu la, to show that the verbal action is carried out toward the deictic center (Mand. fèi lái ‘come flying’, Lahu pò la ‘id.’) (see above 2.3.2).

As a code to quantified expressions lái conveys an approximative notion: wù shì lái su ‘about 50 years old’; liàng mi lái gào ‘about 2 meters high’ (lit. “50 COME years”, “2 meters COME tall”; cf. Eng. give or take 50 years).

So abstract may this verb become that it is frequently used as a pro-verb (much like English do), substituting for verbs of more specific meaning:

(37) nì xiè-xie, ràng wò lái ba
2SG rest let 1SG *** PTCL
‘You rest — let me do (it)’ (e.g. dig, pull, cook, etc.).

(38) wò zài lái ba
1SG self *** PTCL
(a) ‘I’ll do (it) myself.’ (b) ‘I’ll help myself (to food).’
3. VERBLEACHING: FROM CONCRETE VERB TO ABSTRACTER VERB

"Verbleaching" looks as if it should be a German word, but it is actually only a jocular hapology for verb-bleaching, the semantic process by which a full verb undergoes abstractification, at the same time as it expands its syntactic co-occurrence possibilities, or "syntactic versatility". When this process is far enough advanced, the bleached morpheme can become a satellite to an indefinite number of heads — an auxiliary, helping, or "versatile" verb.

3.1. Verb concatenation in Lahu

Even by Lolo-Burmese standards, Lahu is remarkable for the apparent ease with which up to five verbs may be strung together or "concatenated" by simple juxtaposition to form complex verbal nuclei. This is a subject I had to begin to tackle almost a quarter of a century ago (see Matisoff, 1969; 1973: 199–265), and it has acquired an ever fresher interest in the context of the general theory of grammaticalization. Lahu verb concatenation, with its minimum of morphology and its intricate idiosyncrasies, operates according to an interplay of syntactic and semantic factors that resists any rigid formalization. It is in fact a "discourse-driven" aspect of the grammar.

One of the verbs in each concatenation is the verb-head or "Vh". The others (the "juxtapositively productive", or "versatile" verbs) stand in a semantically subordinate relationship to the head, occurring either all to the head's left, or all to its right, or, often, flanking the head on both sides.

Lahu thus accomplishes in a uniform surface way what languages like English need many different structures to do. See Figure 2 (reproduced from Matisoff, 1973: 201), where the verb q3 'hoe' appears in simple juxtaposition with 12 versatile verbs, while the English translations use 6 different types of constructions.

<table>
<thead>
<tr>
<th>Complementary infinitives</th>
<th>q3 ša</th>
<th>'easy to hoe'</th>
<th>qa q3</th>
<th>'help to hoe'</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing complements</td>
<td>q3 ki</td>
<td>'busy hoeing'</td>
<td>tā q3</td>
<td>'start hoeing'</td>
</tr>
<tr>
<td>modal auxiliaries</td>
<td>q3 cz</td>
<td>'should hoe'</td>
<td>gā q3</td>
<td>'must hoe'</td>
</tr>
<tr>
<td>adverbs</td>
<td>q3 bā</td>
<td>'hoe away'</td>
<td>q3 q3</td>
<td>'hoe again'</td>
</tr>
<tr>
<td>prepositional phrases</td>
<td>q3 pi</td>
<td>'hoe for smn'</td>
<td>phō q3</td>
<td>'hoe in a group'</td>
</tr>
<tr>
<td>subordinate clauses</td>
<td>q3 ni</td>
<td>'hoe and see'</td>
<td>ca q3</td>
<td>'go and hoe'</td>
</tr>
</tbody>
</table>

Figure 2: The gamut of English subordinating devices vs. Lahu juxtaposition

The verbs in a true concatenation (there are several other types of multiverbal constructions in Lahu — see 3.13 below) form a single verbal idea, and are all deemed to belong to the same clause. They function as a semantic/syntactic unit.

How my ideas on this subject fell on deaf ears in the late 60's/early 70's! For it flew in the face of a cherished dogma of the generative panjandrums of the day — that each individual verb in a multiverbal "surface" expression must necessarily represent a separate underlying sentence. I felt obliged to defend my monoclusal analysis against the philistines:

It might well be objected at this point that what is actually involved is a nesting of embedded sentences, with each verb in the concatenation deriving ultimately from a separate underlying sentence. For a variety of reasons I find that approach cumbersome for the present purpose. Yet note that it does not really affect the substance of the argument at all whether one operates with embedded sentences or simply embedded verb strings. In either case we are still faced with the problem of stating the complex restrictions on the concatenative process, the nature of the hierarchical relationship among the concatenated entities, etc. (Matisoff, 1973:215)
And how would it account for semantic contrasts between full vs. bleached meanings of the lexeme, according to whether one or more underlying clauses were involved? For sometimes there is a real ambiguity, which proves the correctness of the unitary VP interpretation in one reading. Cf. the Lahu pair of sentences:

(43a) yò yè te chè ve
    3SG house make PROG
    Vh Vv
    'He is building a house.'

(43b) yò yè te (le) chè ve
    3SG house make live
    Vh Vh
    'He built a house and is living (in it).'

(For the criterion of le-insertability in "fortuitous concatenations" like (43b), see below 3.131.)

The non-head verbs in a concatenation get bleached semantically — occasionally in ways that may be idiosyncratic to Lahu, but more often in accordance with well-attested areal (or even universal) tendencies.

Although the grammatical structure of Lahu verb concatenations depends crucially on the semantic features of the individual lexemes involved, brute syntactic constraints also play a role. Though most versatile (= grammatized) verbs follow their Vh’s, about a dozen always precede. We call these two syntactic classes post-head versatile verbs (Vv’s) and pre-head versatile verbs (vV’s), respectively. It is impossible to give a general semantic characterization of the vV’s as opposed to the Vv’s. (One such hypothetical dichotomy would be between “prerequisite” (vV ) and “result” (Vv), but this does not hold water.) This arbitrariness makes it impossible mechanically to assign constituent structure to 3-verb “fore-and-aft” concatenations, which contain both a vV and a Vv in addition to the Vh (see Figure 9, below).

Concatenations containing 2 verbs may thus have either of two hierarchically structures. Either the first verb modifies the second (this is a binary pre-head concatenation, symbolized as vV + Vh), or the second modifies the first (binary post-head concatenations, symbolized as Vh + vV).

3.1.1 Subclasses of post-head versatile verbs.

With only one major exception (see the discussion of ḡa ‘get’, 3.2.2 below) the classes of vV's and Vv's are disjunct. As just mentioned, the set of pre-head versatile differs significantly from the post-head set in terms of number of members: There are only about a dozen vV’s, and they seem to constitute a closed class, while there are several dozen Vv’s, including a number of apparently recent loanwords. The two classes are also strikingly different in their internal ordering properties. The vV’s have few mutual exclusions, and do not fall into semantically coherent subsets. The Vv’s do have such exclusions, making it feasible to divide them into 4 subclasses (i.e. juxtapacapitals, medials, caudals, and variables) on the basis of both semantic and syntactic criteria. (See Figure 3.)

<table>
<thead>
<tr>
<th>Vh</th>
<th>Juxtapacapitals</th>
<th>Medials</th>
<th>Caudals</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vh</td>
<td></td>
<td>Medials</td>
<td>Caudals</td>
<td>Variables</td>
</tr>
</tbody>
</table>

Figure 3: Classes of Lahu Post-head Versatile Verbs

3.1.1.1 The juxtapacapital VV’s

The juxtapacapitals, as their name implies, always occur right next to the head. They number about twelve, and have concrete meanings related to modes of motion or directionality. This is a relatively highly grammaticized group, and the members are mutually exclusive. They are as closely welded to the Vh as are such English “particles” as out or away. (See the examples in Figure 4.)

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING AS Vh</th>
<th>MEANING AS Vv</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka</td>
<td>'put into, insert'</td>
<td>'Vh into'</td>
<td>mā ka 'blow into'</td>
</tr>
<tr>
<td>bā</td>
<td>'throw; discard'</td>
<td>'Vh away; Vh so sthg. is displaced; Vh in an irrevocable way'</td>
<td>tū bā 'burn away'</td>
</tr>
<tr>
<td>yâ?</td>
<td>'descend'</td>
<td>'Vh downward'</td>
<td>pu yâ? 'roll down'</td>
</tr>
<tr>
<td>pō?</td>
<td></td>
<td></td>
<td>yâ? 'jump down'</td>
</tr>
</tbody>
</table>

Figure 4: Some juxtapacapital Vv’s
3.1.1.2 The caudal Vv's.

The caudals occur at or near the tail end of concatenations, and have highly abstract meanings, many connected with potentiality. (See Figure 5.)

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING AS Vh</th>
<th>MEANING AS Vv</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>phë?</td>
<td>'be; be the case'</td>
<td>'able to Vh'</td>
<td>ā phë? 'able to plant'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>qay phë? 'may go'</td>
</tr>
<tr>
<td>gë?</td>
<td>'reach; arrive at'</td>
<td>'manage to Vh'</td>
<td>tu gë? 'manage to stand'</td>
</tr>
<tr>
<td>cë?</td>
<td>'be right; correct'</td>
<td>'ought to Vh'</td>
<td>thë cë? 'ought to touch'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>yë cë? 'should take'</td>
</tr>
</tbody>
</table>

Figure 5: Some caudal Vv's

3.1.1.3 The medial Vv's.

The medials are the most numerous and heterogeneous of the subclasses. They have highly specific meanings (e.g. late, dare, busy, easy) and, like the juxta capitals, are mutually exclusive. This class has a number of adjectival verbs, and is open, since it includes several loanwords from Tai. (See Figure 6.)

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING AS Vh</th>
<th>MEANING AS Vv</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>yë?</td>
<td>'commit an offense; be at fault'</td>
<td>'Vh by mistake; offend by Vh'</td>
<td>dë yë? 'misunderstand'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>së yë? 'lead astray'</td>
</tr>
<tr>
<td>bë?</td>
<td>'be immoderate; be harsh'</td>
<td>'Vh to satiety; Vh to excess'</td>
<td>yë bë? 'sleep one's fill'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>dë bë? 'drink plenty'</td>
</tr>
<tr>
<td>bë?</td>
<td>'be lazy'</td>
<td>'be tired of Vh'</td>
<td>chë bë? 'tired of living'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>gu bë? 'tired of running'</td>
</tr>
</tbody>
</table>

Figure 6: Some medial Vv's

3.1.1.4 The variable Vv's.

The variables are among the most abstract of the Vv's. They have the interesting property of relatively free permutability in "multiversatile" strings (i.e. concatenations with at least 2 versatile verbs), and may occur before or after medials or caudals — with alternative orderings entailing differences in semantic interpretation (see (44) and below 3.13):

(44a) cā phë? chē
cat POT PROG
Vh Vv Vv
'is still able to eat'

(44b) cā chē phë?
eat PROG POT
Vh Vv Vv
'is able to continue eating'

The meanings of the variable Vv's are best characterized as aspectual. (See Figure 7.)

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING AS Vh</th>
<th>MEANING AS Vv</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>chē</td>
<td>dwell/continuative</td>
<td>ni look at/tentative</td>
<td></td>
</tr>
<tr>
<td>ci</td>
<td>send/causative</td>
<td>pā finish/complete; exhaustible</td>
<td></td>
</tr>
<tr>
<td>qay</td>
<td>go/continuative, inchoative</td>
<td>pë give/benefactive; permissive-causative</td>
<td></td>
</tr>
<tr>
<td>mō</td>
<td>be a long time/durative</td>
<td>lā? enough/sufficient</td>
<td></td>
</tr>
</tbody>
</table>

We may illustrate the grammaticized meanings of these verbs by combining them all with the same Vh, e.g. vë? 'put on; wear':

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING AS Vh</th>
<th>MEANING AS Vv</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>vë? chē</td>
<td>'is wearing'</td>
<td>vë? ci 'make/let smn wear'</td>
<td></td>
</tr>
<tr>
<td>vë? qay</td>
<td>'goes on wearing'</td>
<td>vë? mō 'has worn for a long time'</td>
<td></td>
</tr>
<tr>
<td>vë? ni</td>
<td>'wear and see; try on'</td>
<td>vë? pā (1) 'has already put on'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2) 'everybody wears'</td>
</tr>
<tr>
<td>vë? pī</td>
<td>'dress smn'</td>
<td>vë? lā? 'enough to wear; wear enough'</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7: The aspectual nature of the variable Vv's

3.1.2 Pre- vs. post-headedness and the continuum of abstractness.

The reality of the left- vs. right-headedness distinction may be dramatically demonstrated by considering three types of binary concatenations (i.e. concatenations consisting of only 2 verbs, the Vh and a single versatile verb).

3.1.2.1 Where both verbs can function as pre-head versatiles.

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING AS Vh</th>
<th>MEANING AS Vv</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>qhō</td>
<td>(Vh) 'steal'</td>
<td>(vV)'Vh stealthily'</td>
<td></td>
</tr>
<tr>
<td>phō?</td>
<td>(Vh) 'pile up; assemble'</td>
<td>(vV)'Vh in a group'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>qhō phō? (vV + Vh)'assemble secretly'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>phō? qhō (vV + Vh)'steal in a group'</td>
</tr>
</tbody>
</table>
3.1.2.4 The constituent structure and semantic hierarchization of multiversatile concatenations.

Even though there is no correlation between semantic content and preheadedness vs. postheadedness per se, within the pre-head or the post-head group there is definitely a correlation between position in the verb-string and relative abstractness — this can only come out clearly in concatenations of at least 3 verbs (i.e. containing 2 versatile verbs, both of which are on the same side of the head):

\[ V_h + V_v + V_v \quad \text{or} \quad V_v + V_v + V_h \]

Let us here repeat the previous example:

\[(44a) \ cä \ pʰé?\ cä\]
\[
\text{eat} \quad \text{POT} \quad \text{PROG} \quad \text{POT}
\]

\[V_h \ V_v \ V_v \]

'is still able to eat'

\[(44b) \ cä \ cā \ pʰé?\]
\[
\text{eat} \quad \text{POT} \quad \text{POT}
\]

\[V_h \ V_v \ V_v \]

'is able to continue eating'

The general principle is, the closer to the Vh a versatile verb is, the less abstract it is — the more abstract modifies, the more concrete is modified. If all the versatile verbs in a concatenation are on the same side of the head, there is thus no problem in determining the hierarchy of modification. Every outer verb modifies the ensemble of the more inner constituents taken as a unit. Thus, (45) is a pure pre-head concatenation (vC), while (46) is a pure post-head one (Cv):

\[(45) \ ná-ki \ gā \ qā? \ phō? \ lō \ cā \ ve^{15} \]
\[\text{IP} \quad \text{vV} \quad \text{vV} \quad \text{vV} \quad \text{vV} \quad \text{Vh} \quad \text{PTCL} \]

'must again together ask eat'

'We've got to ask (them) again together (for food) to eat.'

The most deeply embedded of the vV's, lō, is subordinate to its head cā (the act of begging is secondary to its envisaged goal, eating; 'beg to eat'. lō cā as a whole is the head of phō?: the 'begging to eat' is performed 'together'. phō? lō cā as a whole is the head of qā?: the action of 'together begging to eat' is performed 'again'. Finally, the string qā? phō? lō cā as a whole is the head of gā: it is the 'repetition of the communal act of

Figure 8: Inherent ambiguity of vV+vV concentrations
supplication to relieve hunger' which is deemed to be 'necessary'.\textsuperscript{15} (See Matisoff, 1973: 214–5)

It can readily be seen that the structure of (46) is the right-branching mirror image of (45):

\begin{align*}
(46) & \quad ci \quad gà\, t5^9 \quad mà \quad pi \quad ci \\
& \quad Vh \quad Vv \quad Vv \quad Vv \quad Vv \\
& \quad \text{tooth pull emerge teach} \quad \text{BENEF ought} \\
& \quad \text{'(They) ought to show them how to pull out teeth.'}
\end{align*}

The futility of trying to write formal rules to capture semantic interpretations is particularly apparent in the case of fore-and-aft concatenations, where there is at least one grammaticized verb on each side of the Vh. Here there is no mechanical way to predict from the syntactic structure what the hierarchy of modification is. Consider the pair of 3-verb concatenations in Figure 9, each of which has the superficially identical constituent structure \(vV + Vh + Vv\):

Those interested in the analysis of maximally complex 5-verb fore-and-aft concatenations like (47) and (48) are urged to consult Matisoff, 1973: 260–4.

\begin{align*}
(47) & \quad gà-hi \quad ga \quad yù \quad qay \quad ci \quad ci \quad ve \quad là \\
& \quad vV \quad Vh \quad Vv \quad Vv \quad VV \quad Q \\
& \quad \text{help-take go} \quad \text{send ought} \\
& \quad \text{‘Should (we) make (them) help to take (it) away?’}
\end{align*}

1. \(ci\) outranks \(ga\) 2. \(ci\) outranks \(ga\) 3. \(ga\) outranks \(qay\) (See Matisoff, 1973: 251, Fig. 27)

\begin{align*}
(48) & \quad gà-hi \quad gà \quad g3^9 \quad chi \quad t5^9 \quad pi \quad ve \\
& \quad VV \quad VV \quad Vh \quad Vv \quad Vv \\
& \quad \text{must return} \quad \text{lift} \quad \text{emerge} \quad \text{give} \\
& \quad \text{‘(We) had to lift (it) out again for (them).’}
\end{align*}

1. \(gà\) outranks \(pt\) 2. \(g3^9\) outranks \(pt\) 3. \(pt\) (automatically) outranks \(t5^9\) (See Matisoff, 1973: 262, Fig. 28)

3.1.3 Other types of multiverbal constructions in Lahu

3.1.3.1 Fortuitous concatenations

Not all Lahu multiverbal sequences represent true versatile concatenations. Sometimes verbs come to be juxtaposed "accidentally", even though...
This particle is an explicit demarcator of the end of a non-final clause, functioning similarly to the “gerundive” or “-TE form” of a Japanese verb:

(49b) tora ga tobidashite kamuitsute kute shimatta to-yu
      tiger PTCL jump out bite onto eat complete QUOT

In a true versatile concatenation it either makes no sense, or else drastically changes the meaning to insert le between the verbs (see, e.g. (43a,b) above). Yet here too the interclausal bonding is a matter of degree. Although the jumping, biting, and eating are 3 successive actions, clearly the latter two verbs, chê? câ are more closely related to each other (almost forming a lexical compound) than either is to pê? ‘jump’. On the other hand, the 4th verb in the string, pê? ‘finish’, is definitely forming a true concatenation with câ, with the grammaticized aspectual meaning ‘completive’ (see above, Figure 7) — so le-insertion is impossible between them. 16

3.1.3.2 Resultative complements

Some resultative verbs are truly versatile, and can occur with grammaticized meanings after an indefinite number of Vh’s. Thus ci ‘stick to; stick fast’ appears both in more “literal” concatenations (e.g. jê? ci (Vh + Vv) ‘stab home (so the knife sticks in the wound)’) and more “figurative” ones (e.g. hâ? ci (Vh + Vv) ‘love unshakably’).

Other resultative verbs are restricted to collocations with only one or two specific Vh’s. Thus tô? ‘catch fire; burn’ occurs only after tô? ‘kindle’; dô? ‘fit inside’ only after kô? ‘insert’ or dô? ‘paek into’; mi? ‘catch’ only after gâ? ‘chase’; kî? ‘be melted’ only after lê? ‘cause to melt’. Verbs of this type, which fail the criterion for “juxtapository productivity”, we call resultative complements (Cr’s). Combinations of Vh + Cr resemble lexical compounds, but differ from the latter in that they must be negated “intrusively”: tô? mô? tô? ‘does not catch fire’ (“light-not-burn”), gâ? mô? mi? ‘cannot catch’ (“chase-not-catch”), etc.

3.1.3.3 Lexical compounds

Lexical compounds composed of 2 free verbs are usually readily distinguishable from versatile concatenations: mô?gâ? ‘smell acrid’ (“stink” + “bitter”), phe-chê? ‘tie up’ (“restrain” + “bind”), chê-mu ‘raise’ (“lift” + “high”). Neither of the elements in a true compound is juxta-productive — each occurs at most in a few combinations with verbs of compatible semantic content. It is as difficult to invent a comprehensible and acceptable Lahu compound as it is to create any neologism. Versatile concatenations, on the other hand, are freely inventable as the communicative need arises, and are not likely to be recognized as novel.

3.1.3.4 Borderline cases.

As always, however, there are borderline cases, where verbal collocations appear to be intermediate between lexical compounds and versatile concatenations. These ‘idiomatic concatenations’ and ‘productive compound-formations’ are discussed briefly in the context of ‘continua of grammatico-semantic categories’ (below 4.3).

3.1.4 Polysemy and grammatization.

Not every meaning shift or glissement sémantique is in the direction of greater abstraction/bleaching/grammatization. Sometimes we just find “lateral slides” in semantic space. Consider the following two Lahu verbs, phê ‘release’ and pô(n) ‘pass by’ (the latter a loan from Tâi).

3.1.4. phê ‘send forth’ > violent action

The verb phê ‘release’ has developed many shades of meaning, only one of which can be said to represent a “grammaticization” — its use as an emphatic Vv, where it lends a vague nuance of outward-thrusting vigor to the verbal action. See Figure 10. 17

3.1.4.2 pô(n) ‘pass by; be spared’ > ago; except

The Tai-derived verb pô(n) ‘pass by; go past’ seems to have originally had spatial/temporal reference. From there it easily passed over into the social (“be spared”) and moral (“transgress”) realms, and finally into the field of abstract logic (“omit in thought”), achieving grammaticized status in a phrase that translates the English conjunction/preposition except. (See Figure 11.)
languages, and which seem to have universal resonances as well: dwell/be at; get/obtain; and give.

3.2.1 'dwell/be at' > continuative/progressive → in/at

Verbs meaning 'dwell; be in/at a place' typically develop into progressive auxiliaries. In some of our SVO languages such morphemes can also function as locative prepositions (see the discussion of 'verpositions', below 4.1).

3.2.1.1 Lahu čhê

a. As main verb:

(50) ᵇᵃ ᵇᵃⁿ ᵏᵃ ᵏᵃ čhê
   1SG gun soul NEG dwell
   'There dwells no soul in my gun.'

b. As post-head versatile verb:

(51) qa-mi čhê šê
   sing *** PTCL
   Vh Vv
   'Go on singing!'

(52) yᵃ-pᵃ čhê giⁿ tō čhê ve tⁿ yⁿ
   boy run roam *** PTCL only PTCL
   Vh Vv Vv
   'This boy just keeps running around.'

As we have noted (above 3.1, [43a,b]), sometimes there is a real ambiguity according to whether the čhê is functioning as an auxiliary or merely as a full verb in fortuitous concatenation:

(43a) yⁿ yⁿ te čhê ve
   3SG house make PROG
   Vh Vv
   'He is building a house.'

(43b) yⁿ yⁿ te (lŠ) čhê ve
   3SG house make live
   Vh Vh
   He built a house and is living (in it).'

3.2 Areal and universal patterns of verbleaching

We cannot attempt here to trace all the interesting semantic developments along the continuum of verb-grammatization in SEA'n languages. Let us pick a few which Lahu shares broadly with other East and Southeast Asian
3.2.1.2 Burmese နိုင်ငံတွင်

a. As main verb:
   (53) ၏ာ၀ိုး နိုင်ငံနှင်
       3SG house be at PTCL
       'He is at home.'

b. As post-head auxiliary verb:
   (54) ၏ားဗားကြည် နိုင်ငံ သို့သာ
       3SG words speak PTCL
       'He is speaking.'

3.2.1.3 Thai ญี่

a. As main verb:
   (55) หงู หม่ำ ญี่ บ้าน
       Hon father NEG be at home
       'Father is not at home.'

b. As post-verbal auxiliary verb:
   (56) หงู เรียน ภาษาไทย ญี่
       3SG study language English PTCL
       'He's still studying English.' (cf. Eng. "He's still at it.")

3.2.1.4 Chinese จี

a. As main verb:
   (57) ตื่น ญี่ บ้าน ญี่ จี
       3SG now be at NEG be at home
       'Is he at home now?'

b. As pre-head auxiliary verb:
   (58) ตื่น ญี่ ญี่
       3SG *** swim
       'She is swimming.'

   (59) ก่อการ ญี่ ญี่
       revolution *** develop people *** advance
       'Revolution is on the ascendant and the people are on the march.'

c. As preposition:
   (60) ตื่น ญา เชิญ ถ้า
       3SG *** house interior read PROG magazine
       'He's reading magazines in the house.'

   (zhe is another progressive marker, grammaticized from a full verb ขอ 'touch'.)

3.2.1.5 Vietnamese 🇻🇳

a. As main verb:
   (61) ông chủ ญา
       HON boss be at home
       'The boss is at home.'

b. As pre-head auxiliary verb:
   (62) ญา ต่อ ญา ญา ญา
       house 1SG *** be near here
       'My house is near here.'

c. As preposition (post- or pre-verbal):
   (1) Post-verbal
   (63) nhiều khi ต่อ ญี่ ญา
       many time 1SG like eat PTCL
       'Often I prefer to eat out.'

   (2) Pre-verbal
   A phrase with ญา may also be preposed to the verb, especially to indicate
   the source of an action:
   (64) ญา ญี่ ญา ญา ญี่
       3SG America go over to
       'come to Vietnam from America'

3.2.1.6 Yao Samsao yim

a. As main verb:
   (65) ญา ญา ญา ญา ญา
       my mother *** DEM CLF house
       'My mother lives in that house.'
meaning ‘have a successful past experience; have managed to Vh; have gotten to Vh’ and/or ‘have to Vh; must Vh’ and in post-verbal position meaning ‘able to Vh’.

3.2.2.1 Lahu ģa

This is the only Lahu versatile verb that can function either as a vV or a Vv.

a. As main verb:

(73) mà ɣa e kāʔ mā bōʔ — ɣa
NEG get PTCL even NEG angry get
e kāʔ mā ha-lè
PTCL even NEG happy
‘If I don’t get it I won’t be mad — if I do get it I won’t be happy.’

b. As pre-head versatile verb:

(1) ‘manage to Vh; get to Vh’

(74) 5-chi me-me ve kāʔ mā ɣa cā ve yō
curry tasty PTCL even NEG *** eat PTCL PTCL
‘We didn’t get to eat very tasty food either.’

(2) ‘must Vh’

(75) chi-bōʔ yā ɣa qay ve yō
now lsg *** go PTCL PTCL
‘I must leave now.’

c. As post-head versatile verb: ‘get by Vh’ing; able to Vh; completed Vh’ing

(76a) jāʔ ɣa ve
stab ***
‘get by impaling’

(76b) Kālā-phu ve 5-chi yā cā mā ɣa
whiteman GEN food lsg eat NEG ***
‘I just can’t eat white men’s food.’

3.2.2.2 Thai dāj

a. As main verb: get/obtain

(77) khōw dāj sataŋ maa cāak khun phōŋ lēśw
3sg obtain money come from HON father already
‘He’s already gotten some money from his father.’
c. As post-head auxiliary (Vv)
In post-head position "tau" (like Thai dâj and Lahu go) indicates ability, possibility, permissibility:

(88) kuv mus tau
   Vh Vv
   'I can go there.'

(89) los tsev puas tau
    come house Q***
    Vh Vv
    'May I come in?'

(90) hsm txiv no noj tau
    fruit CLF DEM eat-able
    Vh Vv
    'This fruit may be eaten.'

In close constituency with a preceding verb, "tau" may indicate the real accomplishment of the purpose of an action (faire et obtenir rééellement quelque chose): nhriav 'look for' > nhriav tau 'find sthg.'; yuav 'buy' > yuav tau 'possess'; caum 'chase' > caum tau 'catch'.

There is a related verb under a different tone, "taus", with meanings quite similar to those of Lahu pi (below 4.41) 'avoir la possibilité physique ou matérielle de; avoir l'habitude de'. Note that ga (mid-tone), the Lahu functional equivalent of Hmong tau, also has a morphotonemically related "alofam" (word-family alternant; see Matisoff 1978) ga (high-falling tone) 'win; overcome'.

3.2.2.5 Vietnamese dârqe and its analogues (especially Thai thûk and Chinese *tak)\(^{18}\)

a. As main verb:
As a main verb, Vietnamese dârqe means (1) 'be acceptable; correct; fine; O.K; all right' (2) 'obtain; get' (3) 'win'.

\[\text{dârqe phêp} \quad \text{obtain authority}\]
\[\text{dârqe múa} \quad \text{get a harvest},\]
\[\text{dârqe trán} \quad \text{win a battle}\]

b. As pre-head auxiliary:
In this position, dârqe is grammaticized into a passivizer (but in the oppositely valued sense from its Thai phonological lookalike, thûk!), conveying the notion of 'being benefited by, affected favorably by some action, state or factor':\(^{19}\)

(91) tòi dârqe dì
    1sg *** go
    vV Vh
    'I was allowed to go.'

(92) tòi dârqe ãn
    VV Vh
    'I am permitted to eat.'

(Contrast (92a): tòi ãn dârqe
    Vh Vv
    'I am edible' (see 3.2.2.5c, below.)

(93) em Toân dârqe thây giaò khen
    y.bro. vv teacher Vh
    'Brother Toan was praised by the teacher.'

(94) mûôî. làm giaò su dârqe chôn di Mô
    10 5 teachers vv Vh go America
    'Fifteen teachers were chosen to go to America.'

c. As post-head auxiliary:
In post-head position, dârqe may indicate ability (in a manner similar to Lahu ga, Thai dâj, Hmong tau):

(95) tòi dì dârqe
    Vh Vv
    'I can go.'

Either verb of (95) may be negated, with a different nuance of meaning:

(96) tòi dì kông dârqe
    Vh NEG Vv
    'I can't go' (because of physical or compelling material reasons).

(97) tòi kông dì dârqe
    NEG Vh Vv
    'I can't go' (because I'm too busy, it's not convenient, etc.).
With some preceding verbs, dārc is a neat translation of the English suffix -able/ible:

ān dārc (“eat + -able”) > ‘eatable, edible’
lām dārc (“do + -able”) > ‘doable, feasible’

As a post-head auxiliary dārc may also have another shade of meaning: That the action of the verb has resulted in something real or good, or has been carried out to a successful conclusion:

(98) mua dārc nhà
buy *** house
Vh Vv
‘bought a house’

(99) sinh dārc mốt đa con-giai
born *** one/a CLF son
Vh Vv
‘had a son born to us’

The position of the object can also affect the semantic interpretation of a post-verbal dārc:

(100a) tôi bát hai con cá dārc
1sg catch two CLF fish ***
Vh Vv
‘I’m able to catch two fish.’

(100b) tôi bát dārc hai con cá
Vh Vv
‘I caught two fish.

Contrast a similar sentence where dārc appears in the pre-head position:

(100c) tôi dārc bát hai con cá
Vh Vv
‘I’m permitted to catch two fish.

Phonologically, dārc resembles the Thai word thūk, and there is a considerable semantic and grammatical overlap between these two important words. The lexical (fully-verbal) meaning of Thai thūk ranges from the moral/abstract sense ‘be right; correct’ to the physical sense of ‘hit the mark; touch; come in contact with.’ There is a homophonous and homographic verb meaning ‘be cheap; inexpensive’ which must also ultimately be related, via the notion the price is right. (See Li Fang Kuei, 1977, p. 102, #16.) As a grammatical morpheme, thūk forms adverbal passive constructions with a following verb or clause: thūk kāt ‘be bitten’, thūk thūrāmaan ‘be tortured’, thūk khāa hāk ‘get one’s leg broken’, thūk rōt chon ‘be hit by a car’. This is in sharp semantic contrast to the Vietnamese ‘benefactive passive’ construction, dārc + Vh (above b).

There is another close phonological resemblance between dārc and Chinese (Old Chinese *tak) ‘get; obtain’, a word with a variety of grammaticized uses in Chinese dialects. This character has several Mandarin pronunciations. As dē it is a full verb ‘get; obtain’. With a different tone and vowel, dēi, it is a modal auxiliary meaning ‘must; have to’, or expressing an unpleasant eventuality. In unstressed form, -de- it marks a positive resultative complement (kām-de-jīn ‘look-can-see’, here being opposed to the negative resultative marker -bu- (kām-bu-jīn” look-not-see’).

The phonological and semantic interrelationships among these Vietnamese, Thai, and Chinese words — and undoubtedly the Yao (Mien) word tū? as well — bespeak a long period of profound language contact.

We may note in passing that there is also a Khmer verb, traw, which fills a range of syntactic/semantic roles quite similar to Thai thūk and Viet. dārc, even though it seems quite unrelated phonologically to the latter two. Thus traw also can mean ‘be correct’ or ‘to hit (the mark)’ as a main verb30, and is also used as a passive complementizer:

(101) pii-msal-mañ kāom traw ckae kham
yesterday 1sg *** dog bite
‘I was bitten by a dog yesterday.’ (Huffman 1970: 302)

3.2.2.6 Khmer baan

Finally, the astounding parallelism in the syntactic/semantic details of the grammatization of words for get in SEA’n languages is also manifested in the Khmer verb baan:

a. As main verb:
(102) look cong baan chaa-khu tee
2sg want *** matches Q
‘Do you want to get some matches?’
b. As pre-head auxiliary:
   In this position, *baan* expresses successful past experience ('to have
   managed to; been able to; to have had a chance to'):

   (103) *kñom baan tiw pii doeng*
   
   1SG  ***  go two time
   
   Vv  Vh
   
   'I've gotten to go twice.'

   (104) *kñom sdp baan*
   
   Vh  Vv
   
   'I can understand.'

   (105) *kñom sdp min baan tee*
   
   Vh  NEG  Vv  PTCL
   
   'I can't understand.'

3.2.2.7 English 'get'

The modal polysemy of *get* comes as no surprise to English speakers. Cf. the ambiguity of (106):

(106) Tomorrow you can get to work.
   a. begin working
   b. have the privilege of working
   c. be transported to work

Sometimes stress helps to disambiguate:

107a) I *gotta [ajgara] swim today. (=I must swim today.)

107b) I *got to [ajgdira] swim today. (=I managed to swim today.)

At any event, the constellation of ideas emanating from the basic notion of 'get; obtain' seems to have much more than areal significance. See the diagram in Figure 12.

Figure 12: Semantic extensions of GET/OBTAIN

3.2.3. 'give' > causative/benefactive

Though many verbs are typically grammatized for causative purposes (e.g. TAKE, DO, MAKE, LET, HAVE, SEND (on an errand)), the benefactive verb *par excellence* is, of course, GIVE.

The Thai causative/benefactive morpheme *kāj* is discussed in connection with postpositions (below 4.2.2). In this section we shall briefly consider some examples from Lahu, Yao, Vietnamese, Khmer, and Mandarin.

3.2.3.1. Lahu *pī and its competitors.

Several verbs have vied for causativizational honors in Lahu, including *yū 'take', te 'do', ci 'send on an errand' (see Matisoff, 1976). The most highly grammatized of them all is *pī 'give':

a. As main verb:
   (108) *phu pī ve 'give money'
   
   OBJ  V

b. As post-head versatile verb:
   (1) benefactive
We have already discussed the paradigmatic opposition between pi (Vv) '3rd person beneficiary' and là (P.), 'non-3rd person beneficiary' in the Lahu VP. (See above 2.32 and Figure 1.)

(109) šu à? tā dē ph-?
3SG PTCL ADV Vh Vv
'Don't scold other people!

(110) yô à? chi qhe qô pi-?
3SG PTCL this like Vh Vv
'Tell that to him.'

causative
In many contexts, the benefactive meaning of pi shades into that of a causativizer (3rd person causee):

ce pi (Vh + Vv) 'make it fall; drop sthg'
chu pi (Vh + Vv) 'make him fat'
ši pi (Vh + Vv) 'make/let him die'

3.2.3.2. Yao Samsoa pun

a. As a main verb:

(111) nin pun pāw yia
3SG Vh axe 1SG
'He gave me an axe.'

b. As a benefactive preposition:

(112) yia tsia? nom daan pun nin
1SG weave CLF basket 3SG
'I wove a basket for him.'

(113) maa càp bûdò-gwây pun fû?-cûay
mother cut fingernails child
'The mother cut the child's nails for him' ('cut the nails for the child').

c. As a causative complementizer (with sentential object):

(114) maa pun fû?-cûay càp bûdò-gwây
mother child cut nails
'The mother let the child cut his nails.'

This sentence is understood as a permissive causative. To express the coercive causative equivalent, a different higher verb, bûa 'tell' must be used, along with a pre-head auxiliary a-tsù? expressing obligation:

(115) maa bûa fû?-cûay a-tsù? càp bûdò-gwây
mother tell child must cut nails
'The mother made the child cut his nails.'

3.2.3.3. Vietnamese cho

a. As a main verb:

(116) ōng áy vêa cho con gài cîêc xe
HON 3SG just daughter CLF car
'He just gave his daughter a car.'

b. As benefactive/causative pre-head auxiliary verb:

cho đân (vV + Vh) 'feed; have someone eat'
cho hay (vV + Vh) 'let someone know; inform smn'

c. As benefactive preposition:

(117) dûa cât cîôi dày cho tòi
hand over CLF broom DET 1SG
'Please hand me the broom.'

d. As permissive-causative complementizer:

(118) ōng áy không cho tòi thôi
HON 3SG NEG 1SG resign
'He wouldn't let me resign.'

e. As benefactive postposition:

(119) đê tòi viêt cho
let 1SG write
'Let me write it for (you).' (Cf. the postpositional use of Thai háj, below 4.22.)

3.2.3.4. Khmer qao

a. As main verb:

(120) manuh prô bân qao sîawphôw taw manuh sîray
person male PAST book to person female
'The man gave the book to the woman.'
b. As a pre-head auxiliary before adjectival verbs:
This is just a special case of the pre-verbal causative effect of qaoy, but this construction translates most naturally by English adverbs (Huffman, 1970: 153): 22

(123) riap-com kluon qaoy chap
prepare self *** quick
‘Get ready quickly.’ (“making it quick”)

(122) sliaq-peeq qaoy sqrt-baat
dress *** careful
‘Dress carefully.’

c. As a causative complementizer (with sentential object):
(123) qaoy maq cong qaoy kNom tswa qa
now mother want *** 1sg do PTCL
‘What would Mother like to have me do now?’

(124) kNom qaoy koot ruti
1sg *** 3sg run
‘I had him run.’ (intentionally)

In this function, qaoy is often used together with the preceding verb tswa ‘do’, to express a more coercive causative. 23

(125) kNom tswa qaoy koot ruti
*** ***
‘I made him run.’ (maybe by scaring him inadvertently)

(126) kNom baan tswa qaoy koot phak sraa klah
PAST *** *** 3sg drink liquor some
‘I made her drink some wine.’

3.2.3.5. Mandarin gēi

a. As main verb:
(127) wōmen yīnggái gēi ta yī ge xīngqī de jià
1pl ought *** 3sg one CLF week GEN vacation
‘We ought to give him a week’s vacation.’

b. As benefactive preposition:
A benefactive prepositional phrase with gēi may either follow or precede the verb of the clause:

(1) Post-verbal prepositional phrase:
(128) xīn yījīng jiǎo gēi ta le
letter already hand over *** 3sg PTCL
‘I’ve already handed the letter to him.’

(129) wō bā yàoshi liù gēi ni
1sg obj key leave *** 2sg
‘I’ll leave the key with you.’

(2) Pre-verbal prepositional phrase:
(130) tā gēi liāo sōng shut dào chā 3sg *** passenger bring water serve tea
‘She brought water and served tea for the passengers.’

Sometimes the meaning of gēi in this construction is more like ‘let; allow’:

(131) gēi wō kàn-kan
*** 1sg look-look
‘Let me have a look!’

If there is a sentential complement, however, gēi cannot be used to convey permission, and the verb ràng (lit. “give way; yield”) is required instead:

(132) tā hài bù kān ràng xiǎohài gén
3sg still NEG willing *** child with
Yōутārēn jiēhūn  Jew marry
‘She’s still not willing to let her child marry a Jew.’

c. As passive complementizer:
In this construction, gēi functions much like the adversative passive marker bèi (see above 3.2.2.5, note 19), following the topicalized patient and introducing the clause which expresses the agency that brought about the unpleasant event:

(133) zhē gū dīrēn quàn gēi yǒuǐ-duí xiāomiè le
this gang enemy all *** guerrilla-band wipe out PTLC
‘The whole horde of enemy soldiers was wiped out by the guerrillas.’

(134) wōmen de yīfù gēi hàn-shut shǐ-tòu le
1pl. GEN clothes *** sweat soak through PTLC
‘Our clothes got soaked with sweat.’
3.2.4. Excursus on Lahu te 'do'

Not surprisingly, te 'do; make' is the Lahu verb with the most abstract meaning and the widest range of grammatical functions. Even in the simplest cases, where it occurs as the only verb in its VP, it collocates with preceding NPs or adverbials in a wide variety of semantic relationships which can only be deduced from the context. To these one may give discrete names (though of course they shade imperceptibly into each other), e.g.: "abstract" (kán te ve "do work"), "factitive" (ā-gě te te 'make a pillow'), "functional" (ghâ-pē te te 'act as headman'), "exploitative" (ā-gě te te 'use a for a pillow'), "appurtenant" (ā-qhê te ve 'have ridges, be corrugated'), "locomotive" (kht-nē te ve "limp"), "adverbial" (gì te ve 'be crooked'), etc.

More abstractly, te may serve as the higher verb governing embedded clauses of various types, e.g. as a "resumptive pro-verb" after two correlative or contrasting verbs:

(135) ẹ̄̄ gị̀ yạ̄̄̄ ẹ̄̄̄ ẹ̄̄̄ ẹ̄̄̄ gị̀ tạ̄̄̄̄ te ve run descend run ascend *** ptcl
'keep running up and down'

This is quite reminiscent of the Japanese construction V₁ -tari V₂ -tari + suru, with the verb suru 'do' used "resumptively" after a pair of correlative verbs, e.g. ittari kitari suru 'be coming and going (all the time)'.

Closer to our present concerns is the use of te as a pre-head versatile verb. Sometimes it retains a more literal meaning in this construction: 'Make something and Vh with it' (te kā ve 'make and put in', lā-gi te te ve 'make tea and set it down'), but often it serves to transitivize or causativize the following verb, whether it be an adjectival (=stative) verb (te qē ve 'widen; make wide'), an intransitive action verb (te pō te 'cause to go through (as a road); cf. French faire pénétrer), or a verb which is already transitive (te tē ve 'cause to crush'; cf. French faire écraser). Most elusively, it sometimes is purely redundant, just adding a bit of phonological bulk but bringing no perceptible increment of meaning. Thus in combination with tạ̄̄̄̄ carry', te has neither factitive nor causative force: te tạ̄̄̄̄ ve can just mean 'carry something along' — it does not necessarily mean either 'make sthg. and carry it', or 'cause to carry'. Similarly, te dē ve just means 'scold smn.' ('do-scold'), not 'cause to scold'.

4.0 CONTINUAM OF GRAMMATICO-SEMANTIC CATEGORIES

There seems to be a fundamental temperamental difference among linguists, according to whether they like neat or messy classifications. In her otherwise perceptive paper on Hmong verb-derived prepositions (1979), Clark comes down squarely on the side of the rigid categorizers:

The question here is, at what point can a word be said to be a preposition and not a verb? The notion of category 'squish' (Ross 1972) or gradual shift from one category to another is as absurd for lexical change as the notion of gradual shift from one sound to another is absurd for sound change. (p. 8)

Somewhat similar is the reductionist position taken in Warotamasikkhadit (1988), who, having made the valid observation that almost all Thai prepositions can be derived historically from full nouns or verbs, goes on to claim that even synchronically there is no such thing as a class of "prepositions" in Thai.

Some grammarians who have dealt with Chinese despair of imposing ready-made "Standard Average European" category-labels on all form-classes and construction types, and have been reduced (or inspired?) to create their own terminology, in a way that some find distasteful. It is as if a proliferation of terminology were some kind of violation of Occam's razor.

Certainly neologistic terminologizing can be carried too far — since linguistic phenomena are continua, it is impossible to find discrete names for every point on each (or any) continuum. One must use heuristic principles — create new terms if necessary, but only if none of the old ones will do. The motive should not be to dazzle or obfuscate by new coining, but rather to clarify.

Yet I personally would like to come down wetly on the side of the squishers. There are duck-billed platypuses in the animal world, and there are borderline entities in grammar as well. I have never shrunken from coining conflationary neologisms to characterize entities which straddle categories: e.g. my Lahu nadverbs, a class of expressions with both nominal and adverbial characteristics (Matisoff, 1973: 306–15). It is far more interesting to set these anomalous phenomena into high relief than to try and sweep them under the rug, or make it look as if they behaved like the non-problematic cases.

4.1 "Verpositions" in SVO languages

In this non-confrontational spirit I would like to propose the euphonious term verpositions for verb-derived morphemes that have come to function like
prepositions. A few random examples:

(a) Viet. vao 'enter' (V) ✸ ‘into NP’ (inessive Prep.)
(136) di sau vao xa-hoi Viet-Nam
‘go deep *** society
‘go deeper into Vietnamese society’

(b) Viet. vo ‘join’ (V) ✸ ‘together with NP’ (comitative Prep.)
(137) tao khoang thich lam voi ong ta
1SG NEG like work *** 3SG
‘I don’t like to work for him.’

(c) Thai khaam ‘cross over’ (V) ✸ ‘across NP’ (transitive Prep.)
(138) phuuuyi dso khaam thanon paj leeew
woman walk **** street go already
‘The woman went off across the street already.’

(d) Mand. yong ‘use’ (V) ✸ ‘with NP’ (instrumental Prep.)
(139) ta yong kuatzi chi bai-cay
3SG *** chopstick eat cabbage
‘He eats cabbage with chopsticks.’

It is interesting to compare and contrast similar serializations in SOV languages. The Lahu translation of (139) is:

(140) yoa y-cu-ka yu le yoa-cay ca ve
3SG chopstick take PTCL cabbage eat PTCL.
(lit. ‘He, taking chopsticks, eats cabbage.’)

Just as the Mandarin serialization with yong ‘use’ is the only way to express the instrumental notion of this sentence, Lahu has no other means than the yu le construction to do so.

But there are several differences here, over and above the relative ordering of verb and object. Chinese can use the yong-construction even with an instrument that is a part of the body:

(141) ta yong shou chi bai-cay
*** hand
‘He eats cabbage with his hands.’

Here Lahu cannot use the yu le construction, but must use the postposed noun-particle ge instead. In Lahu the use of the serialized construction implies volitionality on the part of the agent:

(142) yoa lai-ge ge yoa-cay ca ve
hand Pn

(143) yoa a-tho yu le lai-no tay ca ve
3SG knife take PTCL finger cut PTCL

This can only mean ‘He cut his finger with a knife (on purpose).’ To express the corresponding accidental event the sentence must be recast somehow, e.g.:

(144) yoa ve lai-no a-tho tay ca ve
1SG GEN finger knife cut PTCL
(lit. ‘As for my finger, knife cut (it).’)

The most interesting overt difference between (139) and (140) is the presence of the particle le in the Lahu sentence. The function of this ‘non-final unrestricted particle’ is to indicate that the preceding clause is NOT THE LAST in the sentence. The presence of le makes explicit the biclusal nature of the sentence. (As we have noted (above 3.141), the inserbitvity of le is the chief criterion for distinguishing “fortuitous” from real concatenations.) Yet one cannot help wondering whether the Lahu speaker really considers the verbal event of cabbage-eating to be more of a two-part sequence than the Chinese speaker does. Surely sentence (140) is less biclusal than a sentence like (146):

(146) yoa a-cu-ka a-the le yoa-cay ca ve
bend-break
‘He broke the chopsticks (by bending them) and ate the cabbage.’

There must in fact be a continuum of clausality — a continuum of closeness of relationship or unitarity more subtle than the minimal morphological/syntactic apparatus of these languages can overtly indicate. (Cf. the discussion of the relative clausal autonomy of pairs of verbs in “fortuitous concatenation” above 3.141, ex. (49).)

4.2. Verbs as postpositions in SVO languages

We'll just give a few quick examples:
4.2.1 ‘finish’ > perfective

4.2.1.1 Lahu pà

As a main verb this word means ‘spend; use up’ (phu pà ve ‘spend money’) or ‘finish’:

(147) pà thà qha-gà ʒa te qay ve yò
  finish time until must do go on PTCL PTCL
  Vh Puniv Next vV Vh Vv
  ‘We’ve got to keep on doing it until it’s finished.’

As post-head versatile verb (Vv), pà means ‘finish Vh’ing’ (see above 3.1.2.4):

(148) ʒ cà pà ʒ lā
  rice eat *** PTCL Q
  ‘Have you finished eating?’ (common greeting)

In this post-verbal position pà has developed into a more general perfective marker, in which capacity it can now follow the perfective Pù tā (see 2.3.3 above). Here we should probably consider pà to have become a true verb-particle in its own right:

(149) ʒ tè ni thà ʒà pa pà tā pà
  that one day PTCL 1SG father born PTCL PTCL
  ‘On that day my father was born.’

4.2.2. Thai lèsw

This word does not usually occur as a main verb in Thai, except in some common idioms like pen lèsw kan ‘That’s settled; That does it’, kà lèsw pàj ‘Let it go; Leave it at that’.

In clause-final position this word has been grammaticized into one of the language’s most important aspectual postpositions, signifying ‘completed action; change of state’:

(150) klàp pàj bān lèsw khràp
  return go home *** POLITE
  ‘He’s gone off home already, sir.’

(151) thàa khràydiim mòt lèsw raw ca tham ?oraj
  if beverage used up *** 1PL PTCL do what
  ‘What’ll we do if the drinks are all finished?’

This seems certainly to be an early loan from Chinese (Mand. lào, also a former verb (now literary) meaning ‘to finish’, but grammaticized in unstressed form, la, into a perfective particle. Also undoubtedly due to contact influence from Chinese is the functionally and phonologically similar Hmong aspectual particle lāwm (cf. sentence (87) above; the -m represents a creaky tone in Hmong orthography, while the -w- is barred-i).

4.2.2. GIVE/TAKE ↔ outer-directed/dinner-directed

The Thai verbs hài ‘give’ and ?aw ‘take’ play several interrelated roles in the grammar, where they are more or less neatly, paradigmatically opposed to each other.33

As prepositions, hài indicates intentional benefaction, while ?aw signals that an unintentional effect has been produced on something:

(152) tòŋ jìyun pòon hài phàmākyaan
  must handmoney *** clerk
  ‘You’ve got to hand the money over to the clerk.’

(153) mèw tòkj ?aw sîa
  cat scratch *** mat
  ‘The cat scratched up the mat.’

As complementizers (clause-introducers), hài has a causative function (human or at least animate causee),34 while ?aw has an instrumental one (inanimate instrument):

(154) mèw-khrua hài dêk tāt nỳa pen chèn lêk-lêk
  cook *** child cut meat be slice small-small
  ‘The cook had the child cut the meat into tiny slices.’

(155) mèw-khrua ?aw mûìt tāt nỳa pen chèn lêk-lêk
  *** knife
  ‘The cook used a knife to cut the meat into tiny slices.’

This latter construction is substantially identical to the Lahu instrumental serialization with yù ɂ (above 4.1). We may translate (155) into Lahu as:
be 2-verb sequences composed of the same direction-neutral main verb meaning 'hold', plus either of the pair of grammaticized postpositional verbs meaning come/go to specify the directionality. See Figure 13.

**BRING = HOLD + COME**  **TAKE = HOLD + GO**

<table>
<thead>
<tr>
<th>Language</th>
<th>Verb</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yao</td>
<td>tō</td>
<td>tāy</td>
</tr>
<tr>
<td>Thai</td>
<td>thaw</td>
<td>maa</td>
</tr>
<tr>
<td>Mandarin</td>
<td>mà</td>
<td>lāi</td>
</tr>
<tr>
<td>Japanese</td>
<td>motte</td>
<td>kuru</td>
</tr>
</tbody>
</table>

Figure 13: Bring/take and come/go

### 4.2.3.1. Thai maa 'come' and paj 'go'

(162a) laa  wig khāw paj naj pāa
donkey run enter go into forest
$V_h$ $V_v$ $V_v$ PREP
'The donkey ran off into the forest (e.g. to get away).'

(162b) laa  wig khāw maa naj pāa
donkey run enter come into forest
$V_h$ $V_v$ $V_v$ PREP
'The donkey came running into the forest (e.g. to rejoin his waiting wife and children).'

In this connection we should cite the common Thai greeting, so hard for foreigners properly to understand at first:

(163) paj nāj maa
go where come
'Where are you coming from?; Where have you been?'
('Having gone where do you come?')

### 4.2.3.2. Yao Samsao tāy 'come' and mūng 'go'

The difference in deictic viewpoint exemplified by (162a,b) is exactly paralleled in these Yao sentences:

(164a) bua tō thaw-chāy pāa? kēem tāy
1pl.*** family enter forest ***
'We brought our wives and kids into the forest (to join us there).'
they are most like lexical compounds in that the possible partners of că are restricted to a narrow semantic range.

4.3 Verb-compounds vs. verb concatenations

Certain Lahu verb sequences seem intermediate between lexical compounds and verb concatenations (see Matisoff, 1973, pp. 210–211).

4.3.1. "Idiomatic" concatenations

Occasionally the collocation of a versatile verb with a particular Vh acquires an unpredictable or "idiomatic" meaning. Thus ša (Adj) 'pleasant; comfortable' participates as a Vv in many normal concatenations, where it means 'easy to Vh', e.g. ţa? Ša (Vh + Vv) 'easy to climb', i-kă? qho ša (OBJ + Vh + Vv) 'easy to fetch water', etc. In a few combinations, however, the meaning has become idiomatic:

pī-ša 'be wealthy; prosperous' ("born-easy")
chē-ša 'be healthy' ("dwell-easy")

These are more like lexical compounds than like true concatenations, and it is best to write them with hyphens.

Another interesting example is the compound te-că 'cook' (lit. "do + eat"), which is homophonous with the true concatenation te că (Vh + Vv) 'do for a living' (see 4.32).

4.3.2. Productive compound-formations

The opposite kind of problem is exemplified by the data in Figure 14. The verb că 'eat' appears as an indubitable Vv in collocations where it has the bleached meaning 'Vh for a living'. However, in a number of other collocations with verbs relating to the preparation of food or the killing of an edible animal, că maintains its literal meaning, i.e. 'Vh to eat'. These sequences are like fortuitous concatenations in that the two verbs represent temporally successive actions (though it is not natural to insert le between the elements). They are like versatile concatenations in that the verbs which may precede că in this construction are actually considerable in number. (Lahu happens to abound in verbs relating to the preparation of food.) But

4.4 Versatile verbs vs. verb-particles

4.4.1. Lahu pi 'be able'

The high-frequency morpheme pi 'be able to V; possess the skill to V; possess the quality of Vh to a notable extent' is definitely a verb, since it can be negated (yo mà pi 'cannot speak'). Yet it never seems to occur without another verb preceding, e.g.:

kă pi (Vh + Vv) 'be good at hearing; have sharp ears'
hă? qă pi (Vh + Vv) 'be compassionate' ("able to pity")

This makes pi unique among Lahu posthead versatile verbs, a strange animal less verb than a verb that can constitute a verbal nucleus all by itself, but more verb-like than a verb-particle (which cannot be negated). (See Matisoff, 1973: 679.)
4.4.2. Lahu 'perfective' tá revisited

We have noted (above 2.3.3) how this durative/perfective morpheme seems intermediate between a Vv and a Pv. In *The Grammar of Lahu* I decided (for complicated reasons involving ordering in multiparticle sequences) to consider tá to be a Pv and not a Vv, though I have had misgivings ever since. This problem is aggravated by the fact that as a head verb, tā (whose original meaning is 'place, put, set upright') seems to be conflating with another, accidentally homophonous verb of Tai origin, tā(n) 'set up, establish' (cf. Siamese tăy), in expressions like, e.g., vé-bă? tă ve 'accuse' ('set up a sin').

Curiously, a similar intermediate place on the Vv--Pv continuum is occupied by another perfective morpheme, pó 'finish' (see above 4.2.1.1).

4.5 The syntagmatic continuum: loss of syllable/morpheme boundary

4.5.1. Lahu say 'go' < *qa + e

Certain verbs seem to attract certain particles as collocation-partners — conversely, certain particles occur with particular frequency after a handful of very high frequency verbs. Thus the asseverative particle á is very common after dā? 'be good', or sī 'know' (dā?=á 'Fine! O.K.!', sī á 'I know it (already)!')

The strong declarative particle á is frequent after phē? 'be; be able' (phē? sī 'Sure I can do it!') and chē 'be in a place' (chē á 'Here he is!; It's here all right!).

Just about the only native monomorphemic Lahu word that has the diphthong -ay is the verb say 'go'. Internal reconstruction, as well as comparative/historical evidence strongly indicates that this is an old fusion of a now obsolete verb *qa (< PTB *ka 'divaricate; spread the legs in striding, go') (STC #469) plus e. (Pv) 'transitive motion; motion away from deictic center'. (This is the synonymous counterpart of la (Pv) 'motion toward deictic center' (above 2.3.2).) This particle e must itself once have been a full verb, related to two other words meaning 'go' that survive in limited contexts, viz. i (archaic) as in the ritual phrase sī-mā të phā nā tà i 'Wend not thy way to the land of the dead', and yī (used in set expressions like dā? yī ve 'defecate (poile)'; lit. 'go outside') and yī-yā là-là te ve 'keep going to and fro'). There is also an Akha cognate, i 'go'.

From a broad historical perspective we may thus assume a gradual development from two independent verbs in series, to a sequence of head

verb plus grammaticized directional verb, to verb plus verb-particle, to complete incorporation of the particle into the verb:

V1 + V2 → Vh + Vv → V + Pv → V

4.5.2. Thai khrāp 'polite particle' < khōráp

Equally dramatic is the case of the Thai word khrāp, which is used by men as a sentence-final politeness particle (and as the polite way of saying "Yes" as an independent utterance):

(165) khrāp, khrāp — ca pāj thanthii khrāp

PTCL go right away ***

'Yes, yes — I'll go right away, sir!'

This is historically a fusion of a two-verb sequence, part of the humilistic expression khōráp thōoí 'beg to receive punishment' (khō (V) 'beg', rāp (V) 'receive').

This is quite analogous to the radical reduction undergone by the old English greeting God be with ye! ( > good-bye! > bye!).

4.6. Cheshirization

Finally, we should mention a couple of cases where the phonological reduction of a grammatical morpheme has gone so far that all of its segmental phonemes have disappeared, leaving only a perturbation of the neighboring syllable’s tone as the reminder of its former presence — the smile of the Cheshire cat, gradually fading away.

4.6.1. PTB *sya 'meat; flesh' > *sa- 'body-part prefix' > HIGH TONE

The Lahu root for 'eye' is miţ, under the high-stopped tone. If this syllable had descended from an ordinary (fully voiced) nasal it would have acquired the low-stopped tone. We therefore infer the presence of an *s-prefix at the Proto-Lolo-Burmese level. (This is amply confirmed by forms in the other Loloish languages; see Matisoff, 1972, #145.) The source of this prefix, which can be reconstructed for many other body-part etyme’s, was undoubtedly the independent root *sya 'flesh; animal':

*sya-myak → *samyak → mi? (+ HIGH TONE)
4.6.2. Atsi and Mpi: citation-particle on verb → special tone on verb

Burling (1967) noted the interesting fact that in Atsi (=Zaiwa), a language in the Burnish branch of Lolo-Burmese, words descending from PLB Tone *2 developed different tonal reflexes according to whether they were nouns or verbs. Later a similar phenomenon was observed in the S. Looish language Mpi, where all three of the PLB open tones split according to the form-class of the etymon, yielding a 6-tone system (Srinuan, 1976; Matisoff 1978).

What undoubtedly occurred was the absorption into the verb-root of a citation-particle (actually a nominalizer, like Eng. to in to go). Such a particle is ubiquitous in TB languages, the Lahu form being we (see Matisoff 1972b, and 2.2 above). This particle in many languages is a form with initial vowel — e.g. Akha à — eminently fusible into the syllabic nucleus of the preceding verb. The vowel of the particle must eventually have been swallowed up, leaving only a perturbation of the tone to establish a form-class identity for its head-syllables.

5. Concluding remarks

5.1. Areal and typological differences in grammatizational tendencies.

Although many grammatizational tendencies are doubtless universal, there are certainly areal differences of nuance. Bréal (1897: 16 and passim) points out that Indo-European prepositions usually develop from *adverbs* (an independent and well-defined form-class). In the SVO languages of SE Asia, as we have seen, prepositions derive typically from full verbs.

Brute facts of word-order can be decisive in determining the paths grammatization will take. In the SE Asian linguistic area, verb concatenation has reached its fullest flowering — i.e. strings of up to 5 verbs in a row — in the SOV languages of the Tibeto-Burman family (especially in the phonologically reduced Sinispheric branches, like Lolo-Burmese). These languages are accustomed to bunching up long series of nominal arguments, without verbal intrusions. In their verb-phrases as well they are free to string together long sequences of verbs, without nominal interruptions. On the other hand, the SVO languages (Chinese, Vietnamese, Thai, Hmong-Mien, and the Karetic branch of TB) must break up verb-strings by inserting NPs as close to their governing verbs as possible, so that even their longest concatenations can only achieve a length of 3 verbs (e.g. motion-verb sequences like Chinese fēi jìn qù ‘go flying in’ (see note 39)).

5.2. Predictability and controllability of grammatization

It would seem to be a truism that the course of semantic change is unpredictable in principle, partly because it is at the mercy of extralinguistic factors. We live in a world where a word meaning ‘befigged; stuffed with figs’ (Vulgar Latin *fícātus*) could come to mean ‘liver’ (It. *fegato*, Sp. *hígado*, Fr. *foie*), because of the Mediterranean custom of force-feeding geese with figs in order to cause their livers to hypertrophy into delectable *foie gras*.

The question then arises whether grammatizational meaning shifts are more predictable than semantic shifts in general? Joseph Greenberg has claimed (at the Symposium on Grammaticalization, May, 1988) that this is indeed the case, and that one might safely predict, e.g. that a verb meaning ‘sneeze’ or ‘buy fish’ would never be pressed into grammatical service. While we can certainly agree that such a development is much less likely than for a verb meaning ‘give’ to develop into a benefactive marker, one should perhaps never say *never* when talking about hypothetical semantic developments. Surely in some language or other a verb meaning ‘sneeze’ might acquire a figurative sense like ‘to treat lightly; treat with contempt’ (cf. Eng. *That’s nothing to sneeze at*), giving rise to such grammaticized collocations as ‘to sneeze-talk’ (i.e. ‘talk disparagingly’) or ‘to sneeze-walk’ (i.e. ‘walk insouciantly; strut!’).

Equally unpredictable are the occasional cases of *degrammaticization* that may be observed — instances where an “empty morpheme” has filled itself up and acquired a true lexical root-meaning. Perhaps the best-known example in English is the word for the large conveyance we call *bus*, which is derived ultimately from the Latin dative plural ending of the form *omnibus* ‘for all; for everyone’.

Attempts are sometimes made to intervene directly in the grammaticization process by deliberate language planning. Chinese linguist/administrators have recently taken a paternalistic approach toward “improving” the minority languages of China by encouraging the introduction of new functioners from Chinese, e.g. in order to make inter-clausal semantic relationships more precise. Actually, however, the borrowing or calquing of functioners from more prestigious languages is something that has gone on for millennia even without official encouragement! We have noted the areal diffusion of several functioners above (see e.g. the discussions of Vietnamese *đưới*, Thai *thùk*, Yao *tù* and
Chinese *tok (3.2.2.5); the Thai perfective particle ลืม and Chinese 側 (4.2.1.2), Lahu iversary ‘attitude of regret’ and Thai ต้า (2).

Let us close by trying to predict the fate of two Slavic derivational suffixes in English.

The Russian agentive suffix -nik first impinged on the general American consciousness in 1957, with the successful flight of the first manned satellite, the isskustvennyj sputnik zemli ‘artificial satellite of the earth’, quickly dubbed sputnik (a tri-morphemic form also comprising s- ‘with; together’ and -put-‘journey’; i.e. “fellow-traveler, satellite”). That might have been the end of the saga had not the way already been paved for -nik by words introduced into American English from Yiddish, especially the useful word mudnik ‘pest; pain-in-the-ass’ (from a Slavic root meaning ‘to bother; nag; get on the nerves of’). There is something ingratiating about the sound of -nik, which had long before ensured its success in Yiddish, leading to such made-in-America coinages as allrightnik ‘somebody who’s doing all right for himself; a prosperous person’; neatnik ‘a compulsively neat person’, and nogoodnik ‘scoundrel; sonofabitch’. Yiddish influence was probably cooperating with sputnik to stimulate the brain of the popular columnist for the San Francisco Chronicle, Herb Caen, when he coined the word beatnik around 1960, to refer to members of the “beat generation”. This coinage was spectacularly successful, opening the floodgates for -nik’s further penetration into the lexicon. Soon we had refusenik ‘a Soviet Jew who was refused a visa to emigrate’ — and most recently (February 1989) waitnik, ‘an émigré who is still waiting in Italy for permission to come to the U.S.’ It seems safe to predict that the future holds limitless promise for -nik. It is only a matter of time before such locutions as government-and-bind-nik will catch on.

To this success story we may contrast the bleak future that doubtless awaits another Slavic suffix, even though two words containing it are already familiar to millions of Americans. This is the noun-forming derivational morpheme represented by Polish solidarność ‘solidarity; name of a labor union’ and Russian glasnost’ ‘openness; frankness’. We may confidently venture to predict that this morpheme will never be attached to English roots. In the first place it takes a bit of sophistication to identify the Polish and Russian forms as cognate, so that they do not reinforce each other. Secondly, they are hard to pronounce (especially the Polish variety). Thirdly, they do not sound particularly cute. And finally, they do not fill a derivational or semantic lack — English has plenty of its own nominalizing suffixes, including the phonologically similar -ness.

In general, however, we must content ourselves with sitting back and watching how languages change syntactically and semantically according to their own inner, inscrutable laws. The early generative conception of language change as involving “the addition of rules to the end of a grammar” naively implied that one could completely characterize all the rules of a grammar in the first place. Students of grammaticalization realize that worrying about where one grammar ends and the next grammar begins is a totally meaningless and futile pursuit. For “the new grammar is constantly being created on top of the willing and yielding ruins of the old.”

SYMBOLS AND ABBREVIATIONS

A x B A is an alloform of B; A and B belong to the same word family
CLF classifier
Cr resultative complement
HON honorific
Next extensive noun
Nh head noun; noun-head; head of a NP
Pn noun-particle
Pu unrestricted particle
Puf final unrestricted particle
Punf non-final unrestricted particle
Puniv universal unrestricted particle
Pv verb-particle
Vh head verb; verb-head; head of a VP
Vint intentional verb
Vmod modal verb
vV pre-head versatile verb
Vv post-head versatile verb
PLB Proto-Lolo-Burmese
POT potential
PST Proto-Sino-Tibetan
PTB Proto-Tibeto-Burman
STC Sino-Tibetan: a Conspectus (Benedict/Matisoff, 1972)
WB Written Burmese
WT Written Tibetan

NOTES

1. They eat bugs (as well as wounded members of their own species). The only bad habit they have is laying their droppings on you from the ceiling. This is probably why newly arrived foreigners in Thailand are urged to avoid sleeping with their mouths open.
2. One could point to many similar grammaticizations in the SEA area — e.g. Thai tua ‘body’, tua ‘reg ‘by oneself’ — but the grammaticization of a word for BODY into a pronoun is more a universal than an areal semantic development: cf. Eng. body → indefinite pronouns like somebody, anybody (with reduced stress).

Lahu has a separate word jā-thā-phā for ‘the male body’, which also means ‘male owner’ or ‘he himself’;

(14c) yā jā-thā-phā dā ḫā ‘He has a fine physique.’

3. Incidentally, this expression contains a 4-verb concatenation (see below 3.1), with the main verb po ‘be born’ preceded by two pre-head versatile verbs (ga ‘OBTAiN/must’, ga ‘HELP/auxiliairize’), and followed by the post-head versatile verb pi ‘GIVE/benefactive’.

The various semantic developments of this etymon are diagrammed in a “metastatic flowchart” in Matisoff, 1985: 56.

4. ga may have other shades of meaning as a verb-particle, e.g. ‘like to V; be fond of V’ing’;

5. ga dā-thā-qu cā ga ḫe ‘I like to eat bananas’ (i.e. habitual desire); or ‘imminent likelihood’ (mā-yē lā ḫa ḫe ‘It looks like it will rain’). There is a similar constellation of meanings shared by such particles as Burmese hēʔ, Yao ṣʔyv, and Chrau cərh (above 2.3).

6. The WB cognate la ‘come’ shows that Lahu lā has the etymologically correct tone for the full verb.

7. Verbs meaning ‘come’ (and ‘go’) are eminently grammaticizable, in a variety of syntactic contexts. For a discussion of the Thai postpositions pai ‘GO; away from deictic center’ and maa ‘COME; toward deictic center’, see below 4.1.2. For the use of Chinese lāi ‘come’ as an introduction of purpose-clauses, see below 2.42.

8. In this usage it closely parallels Mandarin lāi appearing as a post-verbal auxiliary: fēi lāi ‘come flying’, pāo lāi ‘come running’, etc.

9. We shall see, however, that in certain contexts it has conflated itself with a distinct Tai-derived morpheme (below 4.4).

10. In this usage it closely parallels the Japanese verb oku ‘put; place’, which has acquired a perfective meaning as an auxiliary (governing the -TE form or gerund of the preceding verb): kangaeru oku ‘think stg. over’, oite oku ‘put stg. down to keep’, totte oku ‘take stg. to keep’, etc.


12. qa? could descend from PLB *kap, while go could be from PLB *kam. (The alternation of homorganic stops and nasals in syllable-final position is one of the most common morphophonemic patterns in Sino-Tibetan word-families.) It is possible that these words are to be related to the etyma reconstructed as PTB *gam (STC #491) ‘put into the mouth; seize with the mouth’ and/or *hup (STC #89) ‘bite; snap at; mouthful’ (both with Chinese cognates). On the other hand, there is a Proto-Tai form *gam ‘mouthful; bite’ (Li Fang Kuei, 1977:198), underlying Thai khom ‘word; speech; classifier for words’. A Tai donor language might well be the immediate source of the Lahu forms. The ultimate relationship of the Proto-Tai and Proto-Sino-Tibetan forms is undoubtedly one of borrowing in one direction or the other, though the resemblance might also be coincidental, involving sound-symbolism (closing the lips at syllable-final to suggest a bite).

13. The benefactive word gěi, here translated ‘for’, is a grammaticalized usage of the verb GIVE (below 3.2.3, 4.2.2).

14. I expanded somewhat on those reasons in a footnote: “It would require a great clanking of generative machinery: obligatory deletions of dummy NP ‘subjects’ (and other ghostly nominal clauses), rules specifying that no adverbs or particles may occur in any of the underlying sentences but the one underlying the Vh, etc. Most crucially, how would it help us to capture the fact that the meanings of the verbs as Vh’s are systematically different from their meanings as versatile verbs?” (Matisoff, 1973: 552, note 43.)

15. To achieve other hierarchical orders of modification, the sentence must be recast, typically by breaking up the concatenation into two pieces, nominalizing one of them and embedding it as the object of the other. (See Matisoff, 1973: 218, Fig. 20.)

16. Despite the ending on kutt, the Japanese verb sequence kutt ebinatt ‘ate up; ate completely’ (lit. “having eaten, finished”), also expresses a unitary monoclusal idea. The -re ending (unlike Lahu le) appears on non-final members of Japanese verb-sequences whether or not the following verb has been “grammatized into” the same clause with it.

17. The theory of these crude semantic diagrams (“metastatic flowcharts”) is developed in Matisoff, 1978. Representations like Figures 10–12 are merely intended to suggest pathways of association in “semantic space”, and are not to be taken in too literal-minded a way. Whatever semantic space is, it is certainly at least non-Euclidean. We can hardly do justice to its unimaginable convolutions by pathetic symbols like arrows and straight lines. A double bar across a line is meant to show that the “boundary of grammaticization” has been crossed, so that on the other side of the bar one is approaching the void of emptied-out, bleached semantic space.

18. The examples in this section are from Thompson, 1965 (pp. 228–9, 309–10, 314, 344–5) and Nguyen Dinh-Hoa, 1966 (p. 147), sometimes with slight changes.

19. Vietnamese does indeed have an adverse passive construction, but this is marked by the auxiliary verb bi ‘suffer; undergo’. This word looks very much like the Mandarin adverasive passive marker bǐ (which originally meant ‘to cover, as with a quilt’), as in:

(90a) Tā hā-ba shì bēi diskū hā-sf de
3SG father COP *** landlord murder FRT
‘His father was murdered by a landlord.’

20. The Khmer expression twe hāy ‘That’s right!’ (“correct already”) corresponds exactly to the Thai phrase thūk hēw.

21. Huffman (1970) observes (p. 75) that in sentences like (103) that include a quantified expression baan may also occur in post-verbal position with little or no meaning difference:

(103a) khom tuv baan pī davō id.’

22. Huffman (1970) transcribes glottal stop by “q”. Some of the examples in this section are from my own field-notes, but they have been more or less regularized into Huffman’s transcription.

23. This is exactly analogous to the causative use of the Thai verb sequence rham + hāj+ Clause (lit. “make + give + Clause”).

24. We is a nominizing particle used in the citation form of verbs or noun-plus-verb collocations; ē is an adverbializing particle.

25. We might dub these camps the neamsik vs. the messnika. For a discussion of our new suffix -nik, see 5.2 below.

26. Cf. Alleton’s (1978) review of Hagège, 1975, where she criticizes such terminological
innovations of the author as adversal (pl. adversals) vs. adversal (pl. adversals), nominatives, predicae, etc. Similar reservations (from an early general category point of view) were expressed by Lehman in his review (1976) of The Grammar of Lahu.

27. E.g., the distinction between Lahu verb-particles (which can only follow verbs) and unrestricted particles (which can follow nouns as well as verbs) is basic to my analytical framework (see above, 2.10). But there are two morphemes which I wish to consider verb-particles (ṣ and ṭ) which can sometimes occur after nouns as well. The pros and cons of the various analytical possibilities are discussed in GL, 339–41.

28. These items have sometimes been called co-verbs in the literature (e.g. Li and Thompson 1974) for Mandarin, Clark (1973) for Vietnamese, Noss (1964: 184–6) recognizes a class of prepositions that are 'weak-stressed forms of verbs'. Parallel examples may be found at will in Hmong-Mien and the Karenic branch of Tibeto-Burman. The most perceptual analysis of Karenic multi-verb constructions yet achieved is to be found in an unpublished dissertation by D.B. Solnit (1983).

29. The first difference lies in the choice of verb each language uses to grammaticalize into the instrumental sense. The Lahu word for 'use' is ɲ (< PLB *sum), which is actually ultimately cognate to Mand. yong (< Old Chin. diung), but while this verb plays no grammatical role in Lahu. On the other hand Chinese does have a co-verb that derives from an old verb meaning 'take', none other than bá, but this has been grammaticalized in a different direction, and is now the highly abstract marker of a pre-verb 'disposed-of' object. It has in fact become the most famous of the co-verbs of the language, with no fewer than 16 papers having been devoted to it during the first 21 years of the annual Sino-Tibetan Conferences!

30. The Pn ge is most often used comitatively, expressing 'with' in the sense of accompany:

   (141a) chi tê há ɲâ-ki ge chê mê
   DEM one night 1PL *** stay PTCL
   'Please stay with us tonight.'

31. This is very similar to the role of the "-TE form" or "gerund" in Japanese verb morphology (see 49b, above). Instrumental NPs are typically marked with the postposition de, as in:

   (145b) nai{xu wo wau-te
         knife *** finger OBJ CUT-PAST
         'He cut (his) finger with a knife.'

This sentence is neutral with respect to intentionality, as is the alternative construction with the gerundive form of tr̂ukau: use:

   (145b) nai{xu wo tr̂ukate yubi wo wau
         OBJ USE -TE

To indicate intentionality, an adverb like waza to 'on purpose' must be added.

32. -the is a verb-combined of the base, resultative type, formed from the verbs ñ 'bend' and the 'break (of an elongated object)'. It may be negated as ñ má the 'not break when bent' ('bend not break').

33. These are cogently discussed in Noss, 1964: 159–62.

34. For a more general discussion of causative/benefactive grammaticizations of GIVE-verbs, see above 3.2.3.

35. Note the use of the feminine agentive nominalizer ɲê-ma (above 2.1.4), nominalizing the verb-compound re-cd 'cook' (below 4.3.1).

36. There is an identical construction in Vietnamese, with cho used as a postposition (see above 3.2.3.3e).

37. The first verb in the Japanese examples are in the gerundive (TE form). See note 31.

38. Cf. the treatment in Noss (p. 185), and Bickner's extensive study (1973) of the discoursal dynamics of these two words.

39. This sort of Thai 3-verb sequence is exactly parallelized in Lahu: verb of specific motion + IN/OUT + COME/GO

   UP/DOWN

   V1 . V2 . V3

   where the first verb expresses the specific manner of locomotion (fly, run, swim), the second signals whether the motion is IN or OUT (Mand. jin vs. chû) or UP or DOWN (Mand. shâng vs. xîâ), and the third tells whether the whole verbal action is performed toward (lái) or away from (qû) the deictic center, e.g. fëi xîâ lâi 'come flying down', pân chû qû 'go running out'.

   The ordering of similar verbs in Yao is somewhat different, e.g.

   Yh-t mâp t 'tshuî (mâp 'go', tshuî 'emerge').

   The latter verb is a loanword from Chinese (cf. Mand. chû).

40. This involves still another continuum, that of etymological identity. At what point do two originally distinct lexemes coalesce into one? Cf. the conflation of English 'let' (as in 'let bald' (tennis)) < Mid.Eng. letten < OE letten 'delay; hinder' and let 'permit' (< ME leten < OE lîten).

41. This verb has itself been grammaticized, with several shades of meaning as a Vv: "go and VH" (lê-gî quy 'go and play'); "Vh along/away (mê quy 'disappear'), 'keep on VH'ing (îe quy 'keep doing'), 'become more and more VH' (chu quy 'get fatter and fatter').

42. As one random example, the high-frequency Hmong conjunction taba 'but' seems to have been borrowed long ago from Chinese dǎn-shî. See C. Li (1988).

43. Matisoff, 1979: 2.

REFERENCES


Frajzynger, Zygmunt. Volume I. "The de dicto domain in language."


Matisoff, James A. 1981. "The bulging monosyllable, or the mora the merrier": echovowel adverbialization in Lahu." in Contributions to Southeast Asian Linguistics;