Phrasal emotion predicates in three languages of Eastern Indonesia

Marian Klamer, Leiden University

In: G.E. Booij, J. van Marle (eds.). 2001. '*Yearbook of Morphology 2000'*.
Dordrecht/Boston/London: Kluwer Academic Publishers, pp. 97-122

Phrasal emotion predicates in three languages of Eastern Indonesia [1]

Marian Klamer, Leiden University

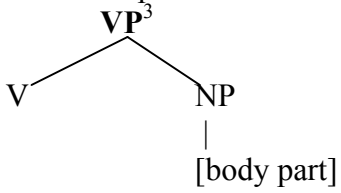
1. Introduction

In many languages emotions are expressed by combining a verb with a body part noun, for example English *My heart bleeds* 'I am sad', and Choctaw *Nok-libisa* 'have a hot neck' > 'be in a passion'.

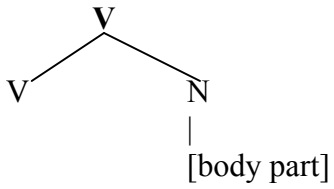
In this paper we examine Verb-Noun (VN) combinations with a similar function in three Austronesian languages of Eastern Indonesia. The languages to be discussed are Tetun,[2] spoken on the island of Timor (Van Klinken 1999), Kambera, spoken on Sumba (Klamer 1998) and Buru, spoken on the island of Buru (Grimes 1991, personal communication, 2000).

In these languages, the VN combinations that express emotions may appear in the syntax both as 'phrasal predicates' and as 'compound verbs'. In a phrasal predicate, V and N are expressed as two separate syntactic constituents, while they constitute one semantic unit and one base for morphological derivations. In a compound verb, N and V are expressed as one (complex) verb, which is a semantic, morphological as well as syntactic unit. The structural distinction is represented in the diagrams in (1). In section 2 I will motivate this distinction.

(1) a. Phrasal predicate



b. Compound verb



The noun in the VN combinations refers to actual body parts, such as 'liver', 'waist', 'head'; to entities related to bodily functions, like 'saliva' or 'breath', or to nouns that express bodily locations, such as 'inside' or 'back'. Apart from describing emotions, the VN combinations also function to describe appearances of character or body. This paper will focus on those that express emotions. Illustrations of the VN emotion predicates are given in (2).

---

[1] This paper and its predecessors have been written at the Vrije Universiteit, Amsterdam, with a fellowship of the Royal Netherlands Academy of Science (KNAW). I would like to thank Geert Booij, Chuck Grimes, Catharina van Klinken for their valuable comments on earlier drafts of this paper. Parts of this paper were presented on the 2000 Pan-Asiatic International Symposium on Languages and Linguistics in Ho Chi Minh City, Vietnam, November 2000, and on the Workshop on Preverbs at Nijmegen University, January 2001. I would like to thank the audiences of both occasions for their input.

[2] The dialect described here is Fehan Tetun, a rural dialect of Tetun spoken in West Timor. It differs greatly from Dili Tetun (Van Klinken, p.c., 2000).

[3] Note that in this conception of Verbal Phrase, the NP is the subject (Theme) of the verb; and is not necessarily its syntactic complement.

(2) Kambera:	V	N		
	hamu	*eti*	‘have a good liver’ > ‘be happy’	
	be.good	liver		
Tetun:	N	V[4]		
	nawan	*sa’e*	‘have ascending breath’ > ‘be angry’	
	breath	ascend		
Buru:	N	V		
	lale-n	*dofo*		
	inside-Poss	be.straight	‘have a straight inside’> ‘be just’	

This paper is structured as follows. In section 2 provide a typological overview of the VN predicates and their expression in Tetun, Kambera and Buru. We will see that the VN predicates occur in continuous and discontinuous syntactic configurations: both as compound verbs and as phrasal predicates consisting of two distinct syntactic phrases (V(P) and NP). Yet, both configurations represent a lexical unit, as I argue in section 3 on the basis of their behaviour in morphological derivations. I propose that not only the compound V but also the VP is a lexically listed unit (cf. Ackermann & Lesourd 1997, Ackermann & Webelhuth 1998, Jackendoff 1997). In section 4 I address the question of how we can account for the mismatch between the syntactic, semantic, and morphological properties of the predicates. In section 5 I propose a scenario for the historical development of the VN emotion predicates that explains the genesis of a new lexical item -- the lexically listed phrase.

2. VN emotion predicates in discontinuous and continuous syntactic configurations

VN emotion predicates in Tetun, Kambera and Buru are expressed in discontinuous and continuous configurations, and these configurations are synchronically coexisting. Illustrations are given in (3) and (4).

(3) Syntactically discontinuous configuration: $_{VP}$[V NP] [5]

a.	Emi	**neon**	keta	**kadolik**
	2p	emotion	don’t	tremble
	‘Don’t (let) your heart tremble’			(Tetun)

[4] The actual order of the verb and the noun happens to be VN in Kambera, NV in Tetun, and is variable in Buru, so the notion ‘VN (emotion) predicates’ as it is used in this paper is mnemonic for “phrasal predicates formed by a combination of a verb and a noun – in any order”.

[5] Glossing conventions for the Kambera, Tetun, and Buru data: ‘1s.(Subj)’, etc. marks the person, number and grammatical function of a pronominal clitic, ‘App’= Applicative, Art= ‘Article’, ‘Attr’= Attributive, ‘Cau’= Causative, ‘Cnj’=Conjunction, ‘Coord’= Coordinator, ‘Dei’= Deictic element, ‘Imm’= Imminent, ‘Irr’= Irrealis, ‘Mod’= Mood, ‘Nom’= Nominaliser, ‘Poss’= Possessive, ‘Prf’= Perfective.

b. **Mbaha** -nanya -ka na **eti**-na na maramba
be.wet -3s.Subj -Prf Art liver -3s.Poss Art king
'The king is pleased' (Kambera)

c. Da **lale** -n **dofo**
3s inside -3sPoss be.straight (Buru)
'S/he is just'

(4) Syntactically <u>continuous</u> configuration: VN predicate is one verb: $_{V}$[VN]

a. Keta **neon** **kadolik**
don't emotion tremble
'Don't (let) your heart tremble' (Tetun)

b. **Mbaha** **eti** -nanya -ka na maramba
be.wet liver -3s.Subj-Prf Art king
'The king is pleased' (Kambera)

c. Da **lal-** **dofo**
3s inside be.straight
'S/he is just' (Buru)

The conceptual structure of the VN emotion predicates minimally includes the three entities in (5), which belong to the word classes in (6). For the sake of concreteness, illustrations from Kambera are given in (7).

(5) a. STATE/EVENT b. THEME c. LOCATION

(6) a. V b. N c. N

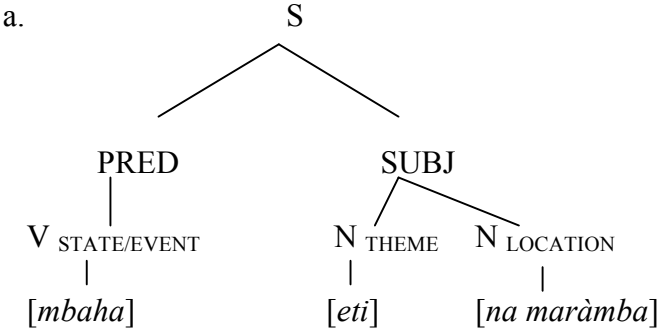
(7) a. *mbaha* 'be wet' b. *eti* 'liver' c. *na maramba* 'the king'

Syntactically, a clause with an emotion predicate such as *mbaha eti* 'have a wet liver' is an intransitive clause – a subject-predicate combination as in (8):

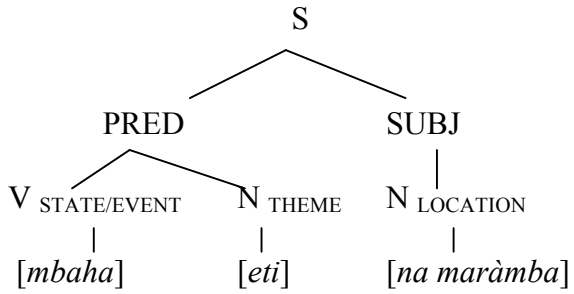
(8) [PRED SUBJ] $_{\text{Clause}}$

When the three lexical items of (5)-(7) are unified with the two syntactic functions in (8), the result can be either one of the two syntactic structures in (9). Structure (9a) is illustrated in (3), structure (9b) is illustrated in (4).

(9) a.



b.



In (9a) we have an emotion predicate whose V and N are expressed discontinuously. The syntactic predicate consists of one verb and the body part noun is expressed in a separate NP. This NP occurs outside the predicate phrase (i.e., the predicate plus the pronominal and aspectual enclitics that attach to it). In the sample sentence (3b), the subject of the clause is *na eti-na na maramba* 'the king's liver'. In other words, the subject of this sentence consists of (i) the Theme argument of the verb (the body part noun), (ii) its Location (the possessor of the body part,[6] which is crossreferenced on the body part noun with a possessive enclitic), as well as (iii) a determiner (the definite singular article *na*). Thus, the subject in (9a) is a full NP, of which the body part noun is the lexical head, whereas its possessor is interpreted as the experiencer of the emotion expressed by the predicate. Note that the syntactic properties of this NP are restricted: it maximally contains a noun, an article and a possessive enclitic; that is, it cannot contain the common range of nominal attributes.

In (9b) we have a complex predicate that is made up of the verb and its Theme (the body part noun). This is a synthetic construction: the predicate is now a compound verb. In such a clause, the only argument that is left to become the grammatical subject is the Location argument -- the possessor of the body part (*na maramba* 'the king').

It is important to observe that the interpretation of (9a) is identical to that of (9b). Both the discontinuous construction and its continuous counterpart are conventionalized metaphors for an emotion where bodily experience is the metaphor of a psychological state ('mind-as-body metaphor', Sweetser 1990: 28-48). That is, the possessor of the body part is interpreted as the experiencer of the emotion in both constructions, and the emotion is expressed by metaphorically combining the verb and the body part noun -- even though these words may belong to different constituents in syntax.

Though the general characteristics of the emotion predicates are quite similar for Tetun, Kambera and Buru, there are also interesting differences between the individual languages. For example, in Kambera and Tetun, one and the same predicate allows for both the discontinuous and the continous construction. But in Tetun, both the continuous and the discontinuous construction are equally allowed, without any apparent differences in the choice of nouns, whereas in Kambera, the continuous construction can only feature the noun *eti*.

The Tetun constructions are illustrated in (10a,b), where the auxiliary *at(u)* can, but need not be positioned between the noun and the verb:

[6] A Location argument in Kambera, Buru and Tetun can be grammatically expressed as either a nominal possessor or as an oblique adjunct. Structural evidence for relating location to nominal possession in Kambera is presented in Klamer (1998:198-199, 1999).

(10) a. Nia at nawan sa'e onan
3s Irr breath ascend Imm
'S/he is about to get angry'

b. Nia nawan at sa'e onan
3s breath Irr ascend Imm
'S/he is about to get angry'

Van Klinken (1997: 199-200, 206-7) reports that all Tetun VN predicates may in principle be expressed as two separate syntactic constituents when they are modified by the auxiliaries and adverbs such as *at(u)* 'Irrealis', *keta* 'don't', *sei* 'still', *hetak* 'increasingly' and *bei* 'also'. In (11b), the negation *keta* intervenes between N and V, in (12b) it is the auxiliary *hetak*:

(11) a. Keta neon kadolik
don't emotion tremble
'Don't (let your) heart tremble'

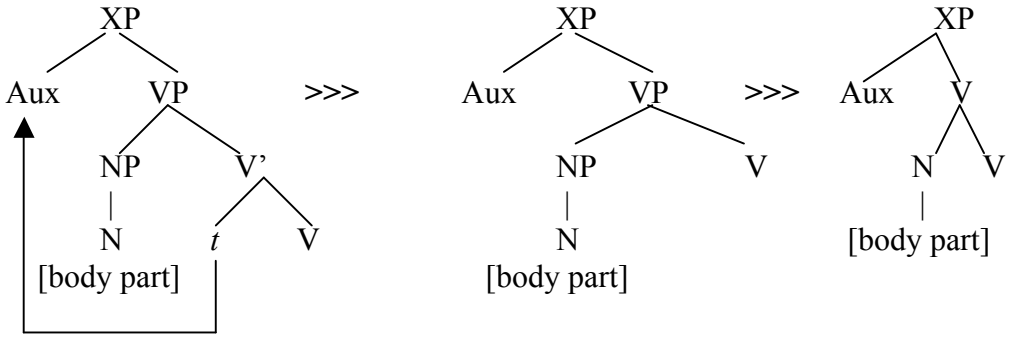
b. Emi neon keta kadolik
2p emotion don't tremble
'Don't (let) your heart tremble'
(reconstructed on the basis of 9.95 & 9.94, Van Klinken 1999:200)

(12) Nia hetak isin kreon
3s increasingly body thin
'S/he grew thinner'

b. Nia isin hetak kreon
3s body increasingly thin
'S/he grew thinner'

Van Klinken (1999:199) also mentions the fact that verbal modifiers directly precede or follow the predicate head in Tetun. That is, the pattern in the (b) sentences of (10)-(12) is the regular pattern for complex predicates in Tetun, and the pattern in the (a) sentences of (10)-(12), where a noun appears between the verbal modifier and the verb, is only possible with emotion predicates. In the configuration in (10a), for instance, the auxiliary *hetak* 'increasingly' modifies *isin kreon* 'thin body', not just *isin* or *kreon*. We therefore analyze the pattern in (10a) as one where the VN predicate is interpreted as a verbal compound. In Tetun, such compounds appear to be syntactically derived: when the combination of a V and a body part N is metaphorically interpreted as an emotion predicate, the auxiliary/adverb can optionally be moved to the left [N [Aux V]] > [Aux [N[V]]]). This movement of Aux is followed by a structural reinterpretation. This reinterpretation is only posssible when N is a body part.

(13) The derivation of emotion compounds in Tetun



In sum, Tetun VN compound verbs are regularly derived from the phrasal construction, with any body part noun, and they are thus the syntactic reflex of the semantic unity of VN emotion predicates.

In Kambera the discontinuous expression is allowed with all types of body part nouns (e.g. waist, mouth, neck), but the continuous construction is only possible with the noun *eti* 'liver'. In other words, (14a) is grammatical, (14b) is not:

(14) a. Mbana -nanya -ka na ngaru-na na maramba
be.hot -3sSubj-Prf Art mouth-3sPoss Art king
'The king is (feeling) malicious'

b. * Mbana ngaru -nanya -ka na maramba
be.hot mouth -3sSubj-Prf Art king

In Buru, too, VN emotion predicates can be expressed discontinuously as well as continuously. In (15a) the V and the N are separate syntactic constituents: both are independent words: syntactically (the N *lale* is marked with a possessive suffix *-n*) as well as prosodically (both *lalen* and *dofo* have main stress). In (15b) the V and the N form a compound: the first word *lale* now has secondary stress and its final vowel is lost. (Main stress is indicated by < ' >, secondary stress by <,>).

(15) a. Da 'lale-n 'dofo
3s inside-3sPoss be.straight
'S/he is just'

b. Da ,lal- 'dofo
3s inside-be.straight
'S/he is just'

The word order in the discontinuous construction is variable. It is unclear which factors determine this. It may be the valency of the base verb: if the verb is transitive, *lale-n* is the grammatical object and usually follows the V, as in (16)-(17). If the verb is intransitive, *lale-n* acts as the grammatical subject, and precedes the V, as in (18), but it may *also*

optionally follow the verb, as in (19) (Grimes, p.c. 2000). (Compare the word order in this sentence with the word order in (15a)).

(16) transitive
Da foni lale-n
3s hide inside-3sPoss
'S/he clams up'

(17) transitive
Da bele-k lale-n
3s be.stupid-App inside-3sPoss
'S/he is confused'

(18) intransitive
Da lale-n boho
3s inside-3sPoss be.bad
'S/he is evil/crazy'

(19) intransitive
Da dofo lale-n
3s be.straight inside-3sPoss
'S/he is just'

In other words, a discontinuous emotion predicate in Buru retains the valency of the base verb and expresses the body part noun as either the grammatical object or the grammatical subject.

The continuous emotion predicates in Buru are much more idiosyncratic than the discontinuous ones. Some examples of such predicates are:

(20)
Da lal- foni-k ii saa
3s inside hide-App some thing
'S/he is keeping something secret (from us)'

(21)
Geba lal- dofo-t
person inside be.straight-Attr
'A just person'

As a rule, the first word in a Buru compound is phonologically reduced: its final vowel is lost and the word cliticizes to the second element, cf. *lale* > *lal* 'inside' in (20)-(21). Buru compounds are (morpho-) syntactically left-headed, so that the category of the first element determines the category of the entire compound. As the first word in the compounds in (15b), (20) and (21) is a noun, they are *nominal* compounds from a morphological point of view. Note, however, that they are interpreted as *verbal* predicates. In other words, compounds such as these, where the order is noun-verb, are exceptional in that not the first, but the second element is interpreted as the head. There are, however, also compound emotion predicates that follow the regular order verb-noun. Examples are (22) and (34b):

(22) Da sus- lale
3s be.difficult inside
'S/he is troubled'

In other words, individual Buru emotion compounds are either VN *or* NV, and their derivational history is unclear. The choice for either order appears to be idiosyncratic, and Buru emotion compounds thus seem conventionalized lexical units. As such, they may

function as the base for further morphological derivations (section 3.4). In contrast to this, discontinuous VN constructions in Buru allow a variation in word order -- at least, when the V is intransitive.[7]

The conclusion of this section is that, though the VN emotion predicates in Kambera, Tetun and Buru are always a semantic unit, in all three languages the discontinuous, phrasal expression of V and N is the most regular and productive construction. This suggests that we need to posit a type of lexical item that consists of a two-word combination and is expressed as a syntactic phrase (or XP). This proposal will be further motivated in sections 3 and 4.

In addition to their regular phrasal expression, the VN predicates may appear as compound verbs, without difference in interpretation. The structural status of the compounds is different for each one of the three languages. In Tetun, the compounds are syntactically derived. In Kambera, emotion predicates are generally expressed analytically, and cannot become a verbal compound, but when the noun is *eti,* verbal compounds *can* be productively derived. Finally, the compound verbs in Buru have so many idiosyncratic features that they must be considered as lexically listed items.[8]

(23) is a summary of the differences between the discontinuous and the continuous construction discussed in this section.

(23) Discontinuous *versus* continuous construction in Tetun, Kambera, Buru

1. *Identical interpretation:*
   Conventionalized metaphor for emotion (mind-as-body metaphor)
2. *Difference in productivity:*
   Discontinuous construction: regular and productive.
   Continuous construction:
   Restricted in Kambera and Buru, syntactically derived in Tetun.
3. *Difference in valency:*
   Discontinuous construction: predicate + two argument positions (Theme, Location).
   Continuous construction: predicate + one argument position (Location).
4. *Difference in thematic content of Subject:*
   Discontinuous construction: SUBJ = Theme
   Continuous construction: SUBJ = Location (> Possessor)

---

[7] Possible discontinuous counterparts of (22) are:

(i)	Da	susa	lale-n	/ Da	lale-n	susa
	3s	be.difficult	inside-3sPoss	/ 3s	inside-3sPoss	be.difficult

'S/he is troubled'

[8] In Buru, a verbal compound can also be derived by incorporating an adjunct nominal (Instrument, Manner, Time, Location) (Grimes 1991:231, 276, 339):

(i)	Da	hai	tu	bohi-n	bika-t
	3s	follow	with	rear-3sPoss	protrude-Nom

'S/he followed with her/his bottom sticking out'

(ii)	Da hai		boh-	biha-k
	3s follow		rear	protrude-App

'S/he followed with her/his bottom sticking out'

It is generally agreed on that the incorporation of adjuncts is a distinctly lexical process, not a syntactic one (cf. the discussion in Spencer 1995).

3. VN emotion predicates are lexical units

3.1. Introduction

Despite the fact that they prefer to be expressed as two separate syntactic units, the VN emotion predicates constitute single lexical units on a number of semantic, syntactic and morphological criteria.

We have seen that semantically, the VN predicates are non-compositional – the verb and the body part noun jointly express *one* emotion, and together have *one* experiencer argument (the possessor of the body part). This interpretation is based on the metaphor that a person's emotional experiences are an event or state of a part of his/her body. Obviously, the metaphor is completely conventionalized – no Kambera speaker would e.g. think that a malicious person literally has a hot mouth, just as no English speaker would think that sadness literally involves a broken heart.

The syntactic fact that the NPs containing the body part nouns cannot include common nominal attributes such as adjectives, numerals, or quantifiers, is a reflection of the non-referential, metaphorical status of the body part nouns. The predicates are also subject to various other types of distributional restrictions. In Kambera, for example, emotion predicates show restrictions in the way the subject can be marked pronominally. Also, the subject NP with the body part noun has a fixed position following the predicate phrase, while in general, subject NPs can also occur in front of the predicate phrase.

Morphologically, the VN combinations are treated as units when they are the bases of morphological derivations. Below we consider the morphological evidence that VN predicates are lexical units, first for Tetun (section 3.2), then for Kambera (section 3.3), and finally for Buru (section 3.4). For the analysis I assume some version of a lexicalist theory of morphological derivations. In a lexicalist theory it is hypothesized that:

(i) Morphological derivations are carried out in the lexicon, not in syntax (the Lexicalist Hypothesis), and

(ii) Syntactic rules neither analyze nor alter word-internal structure (the principle of Lexical Integrity) (cf. Ackerman and LeSourd 1997).

These hypotheses explain the basic and fundamental distinctions between words and affixes. For example, Lexical Integrity accounts for the fact that words are syntactic atoms, while affixes are not. In other words, words can be affected by rules of syntax, as in (24a), while affixes cannot, as illustrated by (24b):

(24) a. This sentence is ungrammatical > Ungrammatical is this sentence
b. * Un- is this sentence grammatical, * Grammatical this sentence is un-

The Lexicalist Hypothesis accounts for the fact that syntactic processes treat both derived and underived words as atoms. This implies, for instance, that syntactic rules do not apply in the lexicon, and that syntactic rules cannot 'look into' the morphological structure of a word. In other words, the lexicalist hypothesis claims that morphological derivation is distinct from syntactic derivation.

Apart from defining the relation between syntax and morphology, lexicalism also assumes specific restrictions on the relation between morphological operations and lexical entries. For example, it is assumed that only *lexical* rules may alter or determine information about the argument structure and valence of a word. Syntactic rules do not do

this. In other words, causative and applicative affixes are part of *lexical* derivational rules, because they change the argument structure of their bases. The prediction is that such lexical rules do not break up morphological objects. (See Ackerman and LeSourd 1997, Ackerman and Webelhuth 1998 and the references cited there.)

In the following sections, I will show that VN predicates in Tetun, Kambera and Buru can be morphologically derived to become causative, applicative, or instrumental verbs, or nominal attributes. Causative and applicative are assumed to be lexical processes because they manipulate the argument structure of the base. Being the bases of such lexical derivations, the VN combinations are also lexical, morphological objects. But we will see that at the same time, the VN combinations do not behave like proper words, because the V and the N constitute separate constituents in syntax. In other words, the VN predicates are lexical units, but they do *not* exhibit lexical integrity and are not syntactic atoms.

3.2. Tetun

In Tetun, the VN predicates can be bases for causativization. Tetun causatives can be periphrastic with the verb *(h)alo* 'make, do', or morphological, with the prefix *ha-*. The VN predicate may be the base for both types of causative (Van Klinken 1999: 199). However, the two constructions have a different word order. In the periphrastic causative, the word order of the base predicate is retained, as in (25a-b), while in the morphological causative it is reversed (NV > VN), as in (26a-b).

(25) a. nawan mohu
breath finished 'be furious'

b. Oan ne'e n-alo ha'u nawan mohu liu
child this 3s-make 1s breath finished further
'This child makes me furious'

(26) a. matan wa' i
eye grow 'wide awake'

b. Ita há hakdiuk hodi ha-wa' i matan
1Pl eat play Coord Cau-grow eye
'We eat snacks to make (us) wide awake'

c. * ..... ha-matan wa'i
Cau-eye grow

These facts are interpreted as follows. The periphrastic construction in (25b) treats the NV predicate as a single, embedded, complex predicate.[9] The causative derivation in (26b), however, treats V and N as separate syntactic constituents (Van Klinken 1999:199; see also p. 84 on compounds). This can be seen by comparing (26b) with the ungrammatical (26c), where the causative has a verbal compound as its base. In other words, the causative in (26b) is a construction where *only* the V is causativized and inflected, while

[9] The predicate can be either continuous or discontinuous in (25b) (Van Klinken, p.c.).

the N (still) functions syntactically as an independent NP (the object). This NP is obligatorily present.

In sum, then, the Tetun causative takes a phrase rather than a word as its base: Causative + $_{VP}$ [V NP]. This is as we expected: in Section 2, we saw that the Tetun VN compound is syntactically, not lexically, derived; therefore we would not expect the compound to be the base of a lexical word formation process like the morphological causative. On the other hand, the fact that the morphological causative takes a phrasal base is evidence that this phrase is indeed available in the lexicon, and is a lexical entry.

3.3. Kambera

Kambera VN predicates, in particular those where N = *eti*, can function as the bases for the derivation of causative, applicative and instrumental verbs. We will first consider the causative and applicative derivations.

Causativization is a very productive process in Kambera. Both intransitive and transitive verbs are transitivized with the causative prefix *pa-*. For example, the stative verb *hàmu* 'be good' becomes *pa-hàmu* 'cause X to be good' > 'improve/restore/relieve X'. In a similar way, the intransitive emotion predicates can also be the bases for causativization. This is illustrated in (27a). Observe that in the causative construction, the noun *eti* must be expressed as a separate (object) NP; (27b) shows that it cannot be incorporated into the predicate. This is the general pattern in Kambera, which does not employ a productive process of noun incorporation (Klamer 1998, chapter 7).

(27) a. Na- pa-hàmu -ya$_i$ [na eti-nggu nyungga]$_i$
3sSubj-Cau-be.good -3sObj Art liver-1sPoss I
'He relieves my heart' (lit.: 'He makes my liver well')

b. * Na- pa-hàmu eti -ya$_i$ [na eti-nggu nyungga]$_i$
3sSubj-Cau-be.good liver -3sObj Art liver-1sPoss I

The indices indicate the crossreference relations between the NP containing *eti* and the pronominal element *-ya* on the predicate. In the normal case, crossreferenced NPs are optional, but when *eti* is involved, the NP is obligatorily present. This indicates that *eti* is part of the (lexical) base of the causative derivation.

Kambera applicatives are derived with the suffix *–ng*, e.g. *pa-hàmu-ng* 'cause (X) to be good for Y' in (28). The nasal suffix is only visible in certain contexts, and for morpho-phonological reasons it disappears when the verb is inflected for its (applicative) object.[10]

(28) Na- pa-hàmu (***eti**) -ngga eti nyungga
3sSubj-Cau-be.good (liver) -1sObj(App) liver I
'He makes me happy/relieved' (lit.: 'He makes my liver well for me')

Again, the noun *eti* cannot be incorporated into the predicate, but must be expressed as a separate and obligatory NP. In other words, though *eti* is an integral part of the

[10] See Klamer 1998, section 6.2, for an account of this alternation.

morphological base of both the causative and applicative derivation, in syntax it is always expressed as a separate constituent.

In the instrumental derivation, on the other hand, V and N are kept together as a compound verb. Kambera instrumental verbs are derived by compounding a transitive or intransitive base verb with the verb *wà(ngu)* 'use', as illustrated in (29). (The final syllable *ngu* is visible in the infinitive form of the verb, but disappears with object marking, cf. (30b)).

(29) palu 'hit X' > palu wa(ngu) 'hit X using Y'
kamakih 'be embarrassed' > kamakih wa(ngu) 'be embarrassed because of Y'

Normally, the object(s) of an instrumental derivation (i.e., the Instrument, and, if the base is transitive, the Theme) are expressed as separate NPs and are not incorporated into the predicate. Example (30) illustrates this for the derivation of instrumental *palu wangu* 'hit X with/using Y': neither the object *tau* 'person' nor the instrument *hurung* 'spoon' can be incorporated into the predicate, cf. (30c,d):

(30) a. Palu wàngu hurung
hit use spoon
'Hit (it) with a spoon'

b. Palu wà-nya$_i$ hurung [na tau nuna]$_i$
hit use-3sObj spoon Art person that.one
'Hit that person with a spoon'

c. * Palu hurung/tau wàngu tau/hurung
hit spoon/person use person/spoon

d. * Palu tau/hurung wà-nya na hurung/tau nuna
hit person/spoon use-3sObj Art spoon/person that.one

The same is true when the base verb is intransitive. In the instrumental derivation of *kamakih* 'be embarrassed', the instrument *ana* 'child(ren)' cannot be incorporated:

(31) Ta- kamakih (*ana) wàngu ana-nda
1pSubj- be.embarrassed child use child-1pPoss
'We are embarrassed about our child(ren)'

Kambera VN emotion predicates with *eti* can be derived to become instrumental predicates:

(32) kudu eti wà(ngu) 'be disappointed with Y'
('have a small liver with/using Y')
bata eti wà(ngu) 'be shattered because of Y'
('have a broken liver with/using Y')
jangga eti wà(ngu) 'be arrogant because/towards Y'
('have a high liver with/using Y')
hàmu eti wà(ngu) 'be happy together with Y'
('have a good liver with/using Y')

karau eti wà(ngu) ‘be angry because of Y’
(‘have a dark liver with/using Y’)

As mentioned above, the rule in Kambera is not to incorporate nouns. In contrast to this, we find that the instrumental derivation of VN predicates expresses the noun *eti* as part of the predicate, as in (33a). It cannot occur as a separate NP, as shown in (33b). This is irrespective of whether the NP containing *eti* is definite and/or possessed, or whether it is only a bare noun; i.e., the presence or absence of the items between brackets in (33b) is irrelevant.

(33) a. Na- jangga eti wà -nda
3sSubj- be.high liver use -1pObj
‘S/he behaves arrogantly towards us’ (lit. ‘S/he has a high liver with us’)

b. * Na- jangga wà -nda (na) eti (-na)
3sSubj- be.high use -1pObj Art liver -3sPoss

In sum, then, though Kambera word formation generally does *not* involve noun incorporation, the instrumental derivation of an emotion predicate with *eti* interprets the V N sequence as a verbal compound.[11] At the same time, however, the causative and applicative derivation of emotion predicates do *not* involve noun incorporation; in such constructions *eti* can only be expressed as an independent syntactic constituent.

The conclusion is that, while the same VN construct is the base for all three derivational processes, the regular syntactic expression of the noun is that of an independent NP. The phrasal construction, not the compound verb with *eti,* is the base for the causative and applicative derivation. Kambera does not productively derive verbal compounds by noun incorporation; and causative, applicative or instrumental derivations are not normally fed by noun incorporation either. Thus, the fact that *eti* is incorporated in the instrumental derivation must be marked as exceptional in the lexicon. In other words, I assume that the instrumental derivation of emotion predicates has the compound verb as its input, while the causative and applicative derivations of the emotion predicates have a phrasal base. Schematically: Instrumental + $_{V}$[V *eti*]; Causative/Applicative + $_{VP}$ [V NP]

3.4. Buru

Buru VN predicates[12] can be the base of causative, applicative, and instrumental derivations, as well as be the base for the ‘attributive’ derivation which derives adjective-like modifiers of nominal elements.

The base of a Buru causative derivation (prefix *pe-/ep-*) is normally a root form like *gosa* ‘be good’ > *pe-gosa* ‘to heal’. Buru VN emotion predicates can also be causativized. The base for the causative derivation may be a VN compound, as in (34), or a phrasal construction, as in (35). The constituent order of the base compound is retained

[11] There is no reason to assume a syntactic incorporation of *eti* for the instrumental derivations, as standard analyses of syntactic incorporation assume that it takes place for reasons of case. In such a scenario we would expect *eti* to incorporate in the applicative derivation as well, contrary to fact.
[12] The data in this section are from Chuck Grimes (personal communication, 2000; Grimes (1991:137-138)).

in the causative derivation. This implies that the causative compound in (34b) has the order verb-noun.

(34) a. ro- (< *roi*) lale
be.small inside 'have a small inside'

b. ep- ro- lale
Cau- be.small inside 'be discouraged, humiliated'

The causative of discontinuous constructions allows more variation, as illustrated in (35b,c):

(35) a. Da dofo lale-n
3s be.straight inside-3sPoss
'S/he is just'

b. Da pe- dofo lale-n
3s Cau- be.straight inside-3sPoss
'S/he reformed her/himself' (lit. 'S/he straightened her/his insides')

c. Da pe- lane-n dofo
3s Cau- inside-3sPoss be.straight
'S/he reformed her/himself' (lit. 'S/he straightened her/his insides')

The instrumental derivation with *lale* always takes the discontinuous construction as its base:

(36) a. Da bele-k lale-n tu ringe
3s be.stupid-App inside-3sPoss with 3s
'S/he is confused with him/it'

b. * Da lal- bele-k tu ringe
3s inside be.stupid-App with 3s

Applicative and attributive derivations take the compound as their base. In (37), the applicative affix *-k* suffixes to the compound. In (38), we find the attributive suffix *-t*.

(37) a. Da foni lale-n
3s hide inside-3sPoss
'S/he clams up'

b. Da lal- foni-k ii saa
3s inside hide-App some thing
'S/he is keeping something secret (from us)'

(38) a. Da lale-n dofo *or* Da lal- dofo
3s inside-3sPoss be.straight 3s inside be.straight
'S/he is just'

b. Ringe geba lal- dofo-**t**

3s person inside be.straight-Attr
'S/he is a just person'

When a VN predicate is the base for both a causative and an applicative derivation in Buru, the base is a compound, as in (39b). But when that verb is put in a sentential context, a final constituent with an additional *lale-n* appears, as in (39c). This constituent acts as the object NP of the causative/applicative verb. Its presence is optional, though preferred (Grimes p.c., 2000).

(39) a. lale gosa
inside be.good 'have a good inside'

b. ep- lal- gosa-k
Cau- inside be.good-App
'cause someone to have a good inside' > 'please someone'

c. Da ep- lal- gosa-k geba di lale-n
3s Cau- inside be.good-App person Dei inside-3Poss
'S/he pleased that person'

The conclusion is that *even* in Buru, with its lexically listed compound predicates, not all the morphological derivations take such compounds as their bases. The applicative and the attributive derivations take a compound verb as their input, the instrumental takes a phrase as its input, and the causative has either a compound *or* a phrase as its base. In other words, also in Buru, we find VN predicates that are lexical units that are expressible as separate constituents in syntax, even after they have undergone morphological derivations.

3.5. Summary

We conclude that in all three languages, the VN emotion predicates are semantically a unit, but are expressible as separate constituents in syntax, even after they have been morphologically derived. In Tetun, the morphological causative of the emotion predicate surfaces is based on a phrase. In Kambera, both the causative and the applicative derivation of the emotion predicates are based on phrases, while the instrumental derivation has a compound verb as its base. In Buru, all of the instrumental, and many of the causative derivations of the emotion predicates are based on phrases, while other causatives, and the applicative and attributive derivation, are based on compounds. A schematic summary is given in (40):

(40) Tetun: Causative + $_{VP}$ [V NP].

Kambera: Causative + $_{VP}$ [V NP]
Applicative + $_{VP}$ [V NP]
Instrumental + $_{V}$[V *eti*]

Buru: Instrumental + $_{VP}$ [V NP]
Causative + $_{VP}$ [V NP], Causative + $_{V}$[V N]

Applicative + $_V$[V N]
Attributive + $_V$[V N]

These data on the derivational morphology of Tetun, Kambera and Buru suggest that the VN emotion predicates in these language are lexical units, even though they do *not* show lexical integrity and are *not* syntactic atoms.

4. How to account for the paradoxical behaviour of the predicates?

We have seen that the VN emotion predicates show paradoxical behaviour with respect to their lexical semantics and syntactic expression in at least four respects:

(41) 1. Morphological derivations of VN emotion predicates can be based on both the compound V and the VP.
2. The VN predicates appear in two distinct syntactic constructions that occur in parallel: the discontinuous construction, which consists of a predicate plus two argument positions (Theme, Location), and the continuous construction, which consists of a predicate and one argument position (Location). In the discontinuous construction, the subject is a Theme argument, in the continuous construction it is a Location argument.
3. The two constructions differ in productivity: the discontinuous construction is regular and productive, the continuous construction is restricted in Kambera and Buru. (It is syntactically derived in Tetun.)
4. There is one interpretation for the two surface appearances of VN predicates.

We have accounted for the first observation by assuming that the lexicon contains items with one *or* two syntactic terminal nodes: compounds ($_V$[VN]) and phrases ($_{VP}$[V NP]). The lexical listing of the VPs accounts for the distributional restrictions on the VP, its semantic non-compositionality, and for the fact that this unit is the base for various types of lexical word formation processes such as causative, applicative, and attributive. I assume that the discontinuous construction is based on the phrasal lexical unit ($_{VP}$[V NP]), while the continuous construction is based on the verbal compound ($_V$[VN]).[13]

The fact that two distinct syntactic configurations have an identical interpretation (observation 4), can be accounted for by assuming that both configurations go back to a single Lexical Conceptual Structure (LCS) (Jackendoff 1997). I propose that the conceptual structure of the VN predicates in Kambera, Tetun and Buru minimally include three semantic entities of the following type (cf. the data in (5)-(7)):

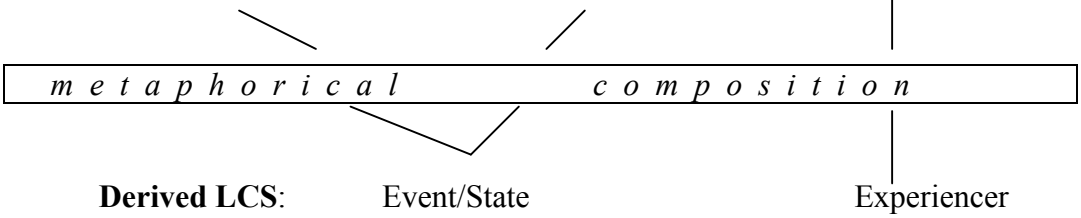
(42) [EVENT/STATE (THEME (LOCATION))]

	[body part N]	[possessor body part]
e.g. ‘be wet’	‘liver’	‘my’

[13] This is true for Kambera and Buru; recall that in Tetun the compound is syntactically derived.

In emotion predicates, that is, the Theme is standardly a body part noun, while the Location is the possessor of that body part. These conceptual entities are subject to the mind-as-body metaphor, where a person's emotional experiences are presented as an event or state of a part of his/her body. This metaphor has been conventionalized and lexicalised. The process involved in this lexicalisation can be seen as a kind of metaphorical semantic composition on the level of LCS, by which a new LCS is derived with two instead of three entities:

(43) **Basic LCS**: Event/State Theme [body part N] Location [possessor]



Derived LCS: Event/State Experiencer

We have seen that the emotion predicates regularly express V and N as distinct clausal constituents. Thus, the Theme and the Location of the basic LCS are projected into syntactic argument positions, and a regular clause is derived, where the Theme becomes the subject of the clause, and the Location is interpreted as the possessor of the subject.

(44)

LCS:	Event/State	Theme	Location
	|	|	|
Syntax:	V	NP_{SUBJ}	NP_{POSS}

Note that in this analysis, the metaphorical interpretation of the emotion predicates is a conceptual process that is not reflected in their syntactic expression. That is, though all the emotion predicates are *interpreted* as a predicate with a single argument (the experiencer of the emotion), this need not be reflected in syntax. In the discontinuous construction, the emotion predicates still have the same number of arguments (two) that it had in the basic LCS.

How does the LCS relate to the two syntactic configurations in which the VN predicates can appear (observation 2)? I assume that the basic LCS projects into the syntactically discontinuous construction, as in (45), while the derived LCS projects into syntactically continuous construction, as in (46):

(45)

LCS:	Event/State	Theme	Location
	|	|	|
Syntax:	V	NP_{SUBJ}	NP_{POSS}

(46)

Derived LCS:	Event/State	Experiencer
	|	|
Syntax:	V	NP_{SUBJ}

The difference in valency between the two constructions is accounted for by assuming that the continuous construction is based on lexically specified compound verbs, which are conceptually derived from the phrasal predicates. The difference in thematic content of the subject is accounted for by assuming that one of the effects of LCS composition is that the Theme argument of the basic LCS incorporates into the derived compound verb, so that the Location is the only argument that remains available to become the subject.

The difference in productivity (observation 3) is explained as follows. The discontinuous construction is the most productive and regular projection of all the VN emotion predicates because this construction represents the most *direct* mapping of LCS unto syntactic structure. The continuous construction is restricted -- it only occurs with lexical compounds -- because this is an *indirect* mapping of LCS unto syntactic structure.

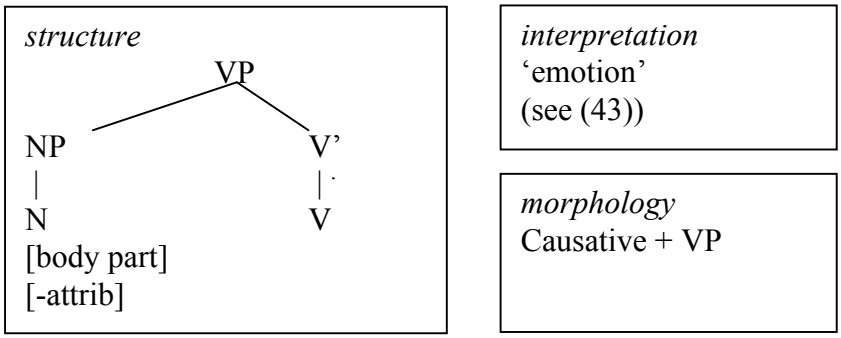
In this analysis, it is assumed that the discontinuous construction is the regular construction, from which the compound is derived. This accounts for the fact that, synchronically, the discontinuous expression is always available, whereas the availability of the continuous construction is more restricted, and varies per language, and even per item: Tetun can syntactically incorporate all body part nouns, Buru appears more selective, and Kambera only incorporates *eti*. The derivation of the compound verb is a lexical derivational process in Kambera and Buru, and the compound is a syntactic atom in these languages. This predicts that the incorporated noun cannot be moved by syntactic rules like topicalization, and that it cannot be modified. It also predicts that a VN compound can be the input for morphological rules. For the Kambera and Buru compounds, these predictions are borne out. Tetun compounds, however, do not function as syntactic islands, because the Aux/Adv can always intervene between V and N. Therefore, the Tetun compounds are analysed as being syntactically derived (cf. (13)). In Tetun, the compounds are not the input for morphological derivations such as causative. Therefore, they are not listed as separate lexical items.

Though Tetun does not have lexically listed emotion predicate *compounds*, the discontinuous emotion predicates in Tetun have features that must be lexically specified somehow. Firstly, because they have a metaphorical interpretation that is not the sum of their parts. Secondly, because the N is unlike other nouns in that it must be a body part noun, and can neither be modified by attributes (nouns, adjectives or verbs), nor moved by syntactic rules such as topicalization (Van Klinken, p.c., 2000). Finally, the discontinuous emotion predicate is the input for certain morphological rules such as causative. In sum, then, the Tetun emotion predicate must also be listed as a separate lexical entry, which has the shape of a phrasal item (VP), as in

(47): [14]

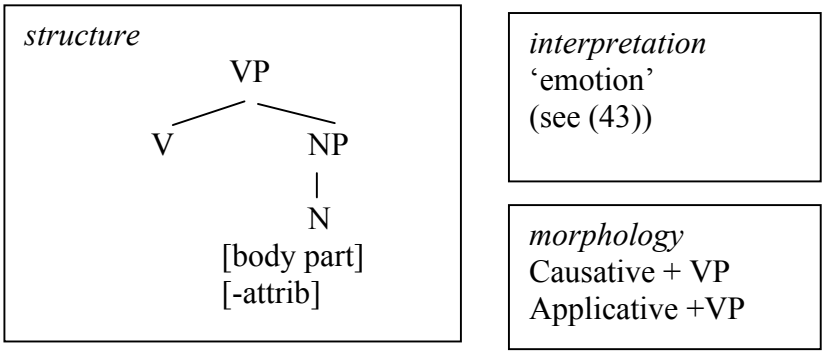
[14] In the following diagrams, the structure of lexical entries is not dealt with in any technical detail. For example, I have collapsed Lexical Syntactic Structure and Lexical Phonological Structure (Jackendoff 1997) in the box '*structure'* and Lexical Conceptual Structure in the box *'interpretation'*. In this paper I am only concerned with pointing out the existence of verbal phrases and verbal compounds as lexical entries. The diagrams are meant as a summary of the various lexical properties of the emotion predicates in the languages at hand.

(47) Tetun: lexical entry for emotion predicate is a VP

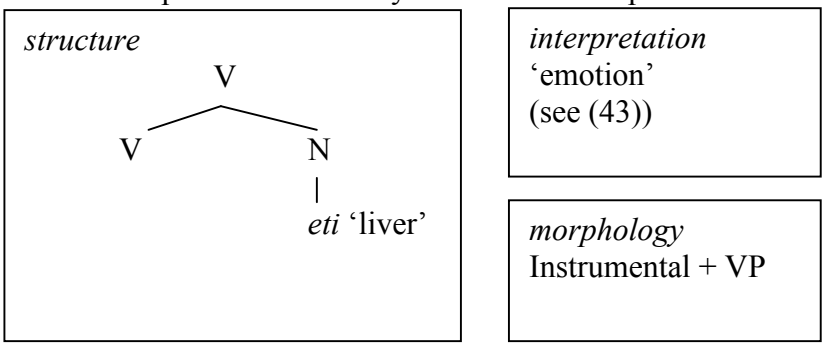


For similar reasons, the lexical entry for emotion predicates in Kambera is also a VP: the VP has a special interpretation, the N must be a body part and cannot be modified by other attributes nor moved by syntactic rules, and the VN predicate can be the input for morphological rules such as causative and applicative derivation. This is represented in (48a). But unlike Tetun, Kambera *also* has lexical compounds to express emotions. The N in these compounds must be *eti* 'liver', the V can be any verb. This compound is the input for the instrumental derivation. It is represented in (48b).

(48) a. Kambera: lexical entry for emotion predicate is a VP

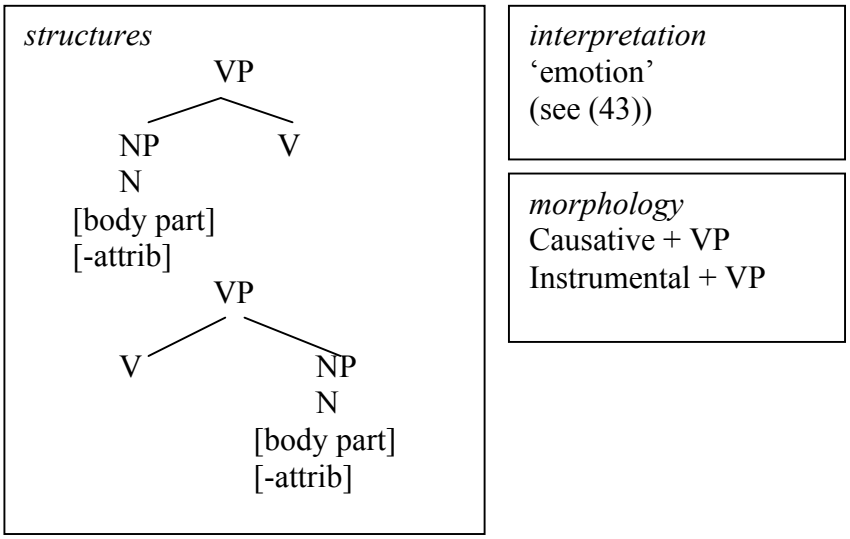


b. Kambera: separate lexical entry for emotion compounds

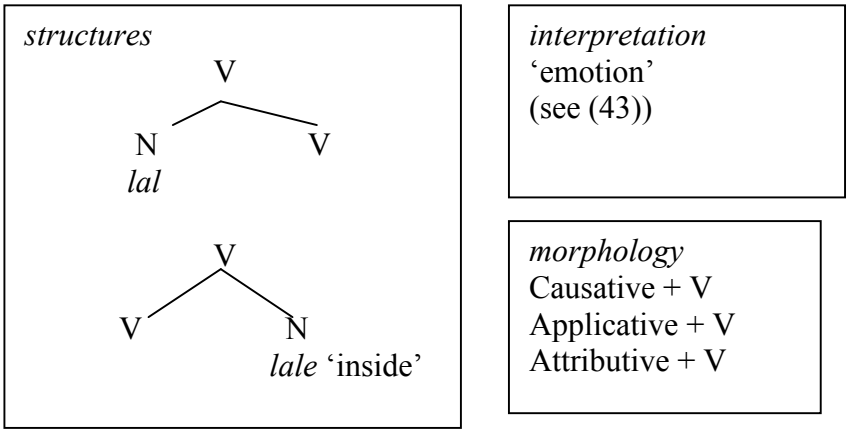


Buru is similar to Kambera: it also has lexically listed VPs and compound Vs for emotions. The word order within VPs and in compounds is variable.

(49) a. Buru: lexical entry for emotion predicate is a VP



b. Buru: lexical entry for emotion predicate is a V



5. The historical development of the VN emotion predicates

In this section I will propose a historical scenario for the development of the VN emotion predicates in Tetun, Kambera and Buru. Given the fact that there are no written records of older stages of these languages, the reconstruction is based on comparative evidence only.

In grammaticalization studies, it is commonly observed that synchronic derivational morphology (including compounding) may be the reflex of a historical change. For the VN predicates at hand, this suggests a scenario where they originated as simple subject-predicate combinations with a metaphorical interpretation of V and N as a single emotion predicate. The metaphorical interpretation became conventionalised and idiomatic. This lead to the lexical listing (in Lexical Syntactic Structure, cf. Jackendoff 1997) of VPs consisting of a V and a body part N and body part nouns. The VP lexical items could subsequently be the base for morphological derivations such as causative and

applicative. Because of the non-compositional interpretation of the VPs, any elements intervening between V and N in syntax were allowed to be moved. The V and the N thus became syntactically adjacent, and were now open to being reinterpreted as compound Vs. But at the same time, the syntactic structure belonging to the original, literal, compositional interpretation of the predicates remained available.[15]

The VN emotion predicates in Tetun, Kambera and Buru witness three distinct grammaticalization stages. Tetun illustrates the initial stage, where the emotion compounds are the result of a syntactic movement. Kambera represents an intermediate stage, where the VN compounds are the result of a productive word formation which is regularly applicable on any V + *eti* 'liver' combination. The process is lexically restricted because it is only allowed with the body part noun *eti.* Buru represents the most lexicalised stage, where all the VN compounds are lexicalised, and the compounds are not (or no longer) transparently derived. A summary is given in (50):

(50) Different grammaticalization stages for the VN emotion predicates

	Discontinuous construction	***Continuous construction***
Stage I: Tetun	Discontinuous syntactic expression of N, V Regular and productive for all body part nouns	Compound verb [N V] $_V$ is syntactically derived Regular and productive for all body part nouns Syntactically derived compounds may become lexicalised
Stage II: Kambera	Discontinuous syntactic expression of V, N Regular and productive for all body part nouns (including N = eti 'liver')	Compound verb [V eti]$_V$ is result of lexical word formation Productive category only when N = eti 'liver'
Stage III: Buru	Discontinuous syntactic expression of N,V or V,N Regular and productive	Lexically listed compound verbs: [lal- V] $_V$ or [V- lale] $_V$ Unproductive but large category

The grammaticalization process discussed in this paper has effects on both the synchronic syntax and the lexicons of Tetun, Kambera and Buru. Syntactically, the languages have two coexisting surface constructions: a discontinuous and a continuous construction. The discontinuous construction is older in origin and synchronically more regular, the continuous construction is relatively new and less regular.[16] Lexically, the

[15] With the observation that the grammaticalization of the VN predicates started off with a semantic reanalysis I do not claim that this is how grammaticalization in general takes place. Elsewhere (Klamer 1999, 2000), I compared the grammaticalization of verbs into complementisers in Kambera, Buru and a third Eastern Indonesian language, Tukang Besi, and I concluded that the change of verbs into complementisers must have started as syntactic reanalysis which resulted in a lexical change, whereby a verb lost an argument. It seems, then, that we cannot generalise about the starting point of grammaticalization. For some phenomena, it is triggered by semantic reinterpretation, for other phenomena, the trigger is syntactic reanalysis. Thus, the question to ask is not: "Does grammaticalization start off with syntactic reanalysis or semantic reinterpretation?", but rather: "*When* does grammaticalisation start off as syntactic reanalysis, and when as semantic reinterpretation?"

[16] Note that the availability of the continuous (morphological) expression of VN by a compound verb does *not* pre-empt the use of the discontinuous/analytic syntactic expression, i.e. there is no morphological 'blocking' involved.

grammaticalization of the VN emotion predicates resulted in the fact that they were lexically listed as phrases (VPs) as well as verbal compounds.

References

Ackerman, Farrell, and Philip LeSourd. (1997). Toward a lexical representation of phrasal predicates. In: Alex Alsina, Joan Bresnan and Peter Sells (eds.). *Complex Predicates*. Stanford, CA: CSLI publications, 67-106.

Ackerman, Farrell, and Gert Webelhuth. (1998).

Grimes, Charles. (1991). *The Buru Language of Eastern Indonesia.* PhD dissertation, Australian National University.

Jackendoff, Ray. (1997). *The Architecture of the Language Faculty*. Cambridge, Mass.: The MIT Press.

Klamer, Marian. (1998). *A Grammar of Kambera*. Berlin/New York: Mouton De Gruyter.

Klamer, Marian. (1999). *The liver as location of emotions.* In: Mathew Pearson (ed.), *Recent Papers in Austronesian linguistics.* UCLA Working Papers in Linguistics.

Klamer, Marian. (2000). 'How report verbs become quote markers and complementisers'. *Lingua* 110, 69-98.

Klinken, Catharina van. 1999. A Grammar of the Fehan dialect of Tetun, an Austronesian Language of West Timor. [Pacific Linguistics C-155]. Canberra: Research School of Pacific Studies, The Australian National University.

Spencer, Andrew. (1995). 'Incorporation in Chuckchee.' *Language* 71, 439-89.

Sweetser, Eve E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural aspects of Semantic Structure.* Cambridge: Cambridge University Press.