National Seminar on POS Annotation for Indian Languages: Issues & Perspectives,12th to 13th Dec, 2011, LDC-IL, CIIL, Mysore.

CLAUSAL GERUND IN MANIPURI: *a combat between the tagger and the tagged*

AMOM NANDARAJ MEETEI

Linguistic Data Consortium for Indian Languages, Central Institute of Indian Languages, Mysore.

CONTENTS

1.0 INTRODUCTION

This paper makes an attempt to concretize the mystifying category of verbal noun in the literature of Parts of Speech Tagging. Verbal noun constructions are very peculiar in that the arguments (subject or object) of a verbal noun can be realized with verbal case marking system such as nominative or ergative or accusative at the clausal level. Following the linguists propounding the syntactic VP projection of verbal nouns (Valoi 1991, Borer 1993, Hazout 1995, Marantz 1997, van Hout & Roeper 1998, Fu, Roeper & Borer 2001, Borer 2005a, 2005b, Park 2008) within the exo-skeleton approach, I also argue that verbal nouns are categorially verbs not nouns and they can be embedded within nominalizing structures in which a derived nominal structure or a gerund structure gets surfaced. Further, this paper explores some of the basic similar properties a clausal gerund behaves in the sense of Pires (1999, 2000, 2001a, 2001b, 2001c) within Minimalist Program approach, dictating that the subject can be either a PRO or an overt DP Case-marked with accusative case or nominative case or ergative case in a class of gerund, hence Clausal Gerund. Finally, this paper suggests for the feature specification of further level of tagging.

2.0 THE STRUCTURE OF VERBAL NOUN

There are three structures (generally two only) that a verbal noun can take. They are given below:

1.1 Derived nominal structure

1.2 Gerund structure

1.3 Clausal gerund structure

Derived nominal vs Gerund structure:

When a verbal noun takes the transitive argument structures consisting of an agent and a theme argument, the structure in which the theme argument is genitive-marked is the Derived Nominal one; whereas, the structure in which the theme argument is accusative-marked is that of Gerund one. This case is illustrated below:

GEN ACC

(1) a. yeknaba-na [$_{DP}$khungang-**gi** mAngnaba]-bu taukhi
enemy-ERG village-GEN destruction-ACC did
'The enemy destroyed the village'

ACC ACC

b. yeknaba-na [$_{DP}$khungang-**bu** mAngnaba]-bu taukhi
enemy-ERG village-ACC destruction-ACC did
'The enemy destroyed the village'

2.0 The Structure…

- What is shown here is that the sentence (1a), where a theme argument is genitive-marked, instantiates the case in which a verbal noun takes a derived nominal structure while the sentence (1b) with an accusative-marked theme illustrates the case where a verbal noun takes a gerund structure.
- Numerous linguists (Chomsky 1970, Abney 1987, Grimshaw 1980, Valoi 1991, Harley & Noyer 1997, Borer 1999, Alexiadou 2001) have stipulated contrastive properties of derived nominals and gerunds.

  a. Derived nominals exhibit properties similar to a typical NP and they can take adjectival modification, but they do no have the ability of verbal case marking.

  b. Gerunds have the properties of VP and they cannot take adjectival modification but take adverbial modification. They can assign accusative case to an object if present.

The structure…

The Clausal Gerund

It is a class of gerund, in which the subject can be either a PRO or an overt DP Cased-marked with accusative Case (acc-ing) or with nominative Case (Pires 2006:15). . Let us briefly see the following examples:

a. Jack worried about ***PRO being*** late for dinner

b. Jack worried about ***John/him being*** late for dinner.

We now see that there are two gerund structures that in (1a) above, the subject is PRO (TP-defective gerund) and, (1b) above, the subject is lexical. There is no alternation between PRO and overt subject in either type of structure. Let us see the following Manipuri examples:

2. a. Tomba-na ***PRO*** cA-ba pAm-de
   Tomba-ERG eat-NMLZ like-NEG
   Tomba does not like (PRO) to eat.

   b. Tomba-na ***mA-bu*** cA-ba pAm-de
   Tomba-ERG he-ACC eat-NMLZ eat-NEG
   Tomba does not like him to eat.

In the similar fashion, the verbal noun *cA-ba* 'eating/to eat', realizes the subjects as PRO or overt accusative-marked DP subject ***mA-bu 'He-ACC'***.

3.0 VERBAL NOUN AS A VERB FOR POS TAGGING

There are two possible options as far as the grammatical category of a verbal noun is concerned:

(3) a. Verbal Noun as Noun, as the term 'verbal noun' advocates,

b. Verbal Noun as Verb, as suggested from the morpho-syntactic properties .

(3)a. Verbal Noun as Noun:

Many linguists (Baker 1988, Ahn, Kageyama 1991, Miyamoto 1999) consider the syntactic incorporation account, suggesting that verbal nouns as nouns incorporate into a light verb at the level of syntax. Let us see the example (3) below (example 1b above):

(4) yeknaba-na [khungang-**bu** [mAngnaba taukhi]

enemy-ERG village-ACC destruction did

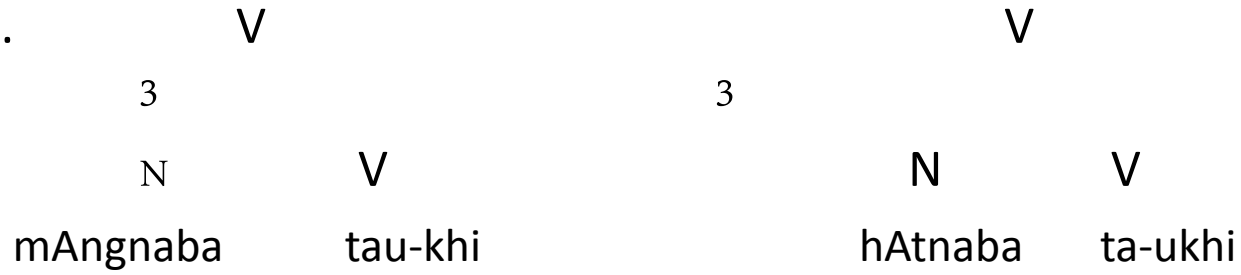
'The enemy destroyed the village'

Under this analysis, verbal noun mAngnaba 'destruction' raises to a verb head tau 'do' at the level of syntax and adjoins to it, which results a complex predicate formation. But, the arguments of verbal noun become free to raise up to the verbal domain to get case.

VERBAL NOUN AS VERB…

Another approach (Miyagawa 1989) suggests that incorporation whereby a verbal noun bonds with a light verb takes place in the lexicon, as given in the structure 5(a) & (b) below:

5a.



It says that the complex predicate s as one word are inserted into a syntactic structure under a V node, predicting that the N-part corresponding to a verbal noun would not be syntactically visible, since the predicates are found derived in the lexicon and thus inserted under V nodes in syntactic structures.

(3)b. Verbal Noun as Verb:

I follow the linguists propounding the syntactic VP projection of verbal nouns (Valoi 1991, Borer 1993, Hazout 1995, Marantz 1997, van Hout & Roeper 1998, Fu, Roeper & Borer 2001, Borer 2005a, 2005b, Park 2008) within the exo-skeleton approach, and also argue that verbal nouns are categorially verbs not nouns and they can be embedded within nominalizing structures in which a derived nominal structure or a gerund structure gets surfaced. Following are some of the main factors:

(i) Adverbial modification (ii) Verbal Noun Stacking (iii) Constituent Structures

VERBAL NOUN AS VERB…

(i) Adverbial Modification:

According to Baker 1983 et., the verbal noun part is not syntactically visible within the complex predicate. Let us see the examples 6(a) & (b) below:

6 a. yeknaba-na konung-du-bu loyna koysinba ngam-kha-re
enemy-ERG fort-DST-ACC completely round can-CERT-PERF
'The enemy could round the fort completely'

b.* yeknaba-na konung-du-gi koysinba ngam-kha-re
enemy-ERG fort-DST-GEN completely round can-CERT-PERF
'The enemy could round the fort .'

In 6(a) above, the verbal noun can't take an adjective, and instead, it takes an adverb. And, the syntactic incorporation account predicts that the verbal noun, as a noun, can take a genitive argument as its complement, but the finding fact fact is contradictory to the prediction as shown in 6(b) above. This shows that the verbal noun part of complex predicate is syntactically not visible, and hence the assumption that verbal nouns are nouns is incorrect.

Again, regarding Miyagawa (1989)'s account of lexical derivation in the lexicon, let us see the following examples (7) & (8):

VERBAL NOUN AS VERB…

(7). lAlmising-na thong-bu yAnkhaiba tau-khi, aduga yum-di tau-khi-de.
soldiers-ERG bridge-ACC breaking do-CERT, but house-TOP do-CERT-NEG
'The soldiers broke the bridges but did not the house.'

(8). a. lAlmising-na thong-bu yAnkhaiba tau-kha-ra-ba-ra?
soldier-PL-ERG bridge-ACC breaking do-CERT-PROS-NMLZ-INT
'Did the soldiers break the bridge?'

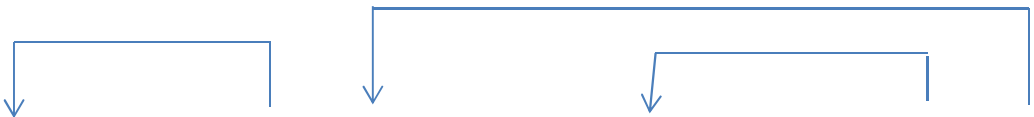
b. hoi, Φ_i tou-kha-re.
Yes do-CERT-PERF
'Yes, they did.'

We now see that only the verbal noun part in complex predicates can undergo the syntactic operation of ellipsis to the exclusion of a light verb, which is in contrast to the prediction of lexical incorporation account. Hence both the syntactic incorporation and lexical derivation analyses are inconsistent with the Lexical Integrity Hypothesis(LIH) (Lapoite 1979), stating that the internal structure of a word cannot be relevant in syntax. Hence, verbal nouns are actually verbs and they take their arguments simply because they are verbs (Park 2008).

VERBAL NOUN AS VERB…

(ii) Verbal Noun Stacking:

As an evidence for the existence of syntactic VP, verbal nouns also exhibit verbal properties such as assigning accusative case to their arguments and licensing adverbial modification. It so happens when one verbal noun follows another verbal noun, i.e., verbal noun stacking, a bare verbal noun shows the ability to assign accusative Case similar to a verb, as shown in (9) below:



(9) Hajari-na [[karapsan-**bu** thijinba]-**bu** **mapungphAna** saugatpa]-bu
Hajari-ERG corruption-ACC investigation-ACC completely supporting-ACC
thagatpa-bu taukhi
thanking-ACC did

In (9) above, the verbal noun *thijinba* 'investigation' assigns accusative case on karapsan 'corruption'. The verbal noun *saugatpa* 'supporting' licenses the modification by the adverbial *mapungphAna* 'completely'. Since there is no intervening light verbs to support the verbal nouns to take verbal properties and it obeys the Head-to-Head movement constraint (HMC), it signals the presence of a syntactic VP element.

VERBAL NOUN AS VERB…

(iii) Constituent Structures:

Following examples show that verbal nouns can be explained as derived nominals or gerunds.

< Topicalization >

(10) a. *[mAngnaba]-di yeknaba-na khungang-gi taukhi
destruction-Top enemy-ERG village-GEN did

b. [khungang-**gi** mAngnaba]-di yeknaba-na taukhi
village-GEN destruction-Top enemy-ERG did

(11) a. *[mAngnaba]-di yeknaba-na khungang-bu taukhi
destruction-Top enemy-ERG village-ACC did

b. [khungang-**bu** mAngnaba]-di yeknaba-na taukhi
village-ACC destruction-Top enemy-ERG did

< Scrambling >

(12) a. *[mAngnaba]-bu yeknaba-na khungang-gi taukhi
destruction-ACC enemy-ERG village-GEN did

b. [khungang-**gi** mAngnaba]-bu yeknaba-na taukhi
village-GEN destruction-ACC enemy-ERG did

VERBAL NOUN AS VERB…

(13) a. *[mAngnaba]-bu yeknaba-na khungang-bu taukhi
destruction-ACC enemy-ERG village-ACC did

b. [khungang-**bu** mAngnaba]-bu yeknaba-na taukhi
village-ACC destruction-ACC enemy-ERG did

I follow Park (2008) in that the theme argument *khungang* 'village' and the verbal noun *mAngnaba* 'destruction' form one single constituent DP. Since movement should observe a constituent structure, the ungrammaticality of each (a) sentence obtains a straightforward account. Hence, each (b) sentence should be grammatical as it observes a constituent structure. This shows that verbal noun phrases can analyzed as derived nominals or gerunds and such prediction is done through the movement operations such as topicalization or scrambling.

4.0 CLAUSAL GERUND (CG) STRUCTURE

4.1 Some properties of CGs:

Pires (2006) proposed the analysis of the syntax of CGs attempting to account for five core syntactic properties of clausal gerunds, regarding especially their distribution and licensing of subjects within Case checking/valuation approach under the Minimalist program (Chomsky 2000, 2001).

i). The subject of a CG may be an empty category (standardly analyzed as a PRO) or an overt DP:

English:

(14) a. Jack worried about ***PRO being*** late for dinner

b. Jack worried about ***John/him being*** late for dinner.

Manipuri:

(15) a. Tomba-na ***PRO*** cA-ba pAm-de
Tomba-ERG eat-NMLZ like-NEG
Tomba does not like (PRO) to eat.

b. Tomba-na ***mA-bu*** cA-ba pAm-de
Tomba-ERG he-ACC eat-NMLZ eat-NEG
Tomba does not like him to eat.

PROPERTIES OF CGs…

ii). CGs need to satisfy a Case requirement:

English:

(16) a. *Mary talked about [(that) John moved out]

b. Mary talked about [John moving out]

Manipuri:

(17) a. *Tomba-bu$_i$ ai-na [t$_i$ catpa] pammi

Tomba-ACC I-ERG going like

b. ai-na [Tomba-bu catpa] pammi

I-ERG Tomba-ACC going like

iii). CGs do not behave as Exceptional Case Marking (ECM) complements:

English:

(18) a. Mary believes [Paul to be smart]

b. *Mary believes [John being smart]

Manipuri:

(19) a. Tomba-na Ibemma-bu phaja-i thAja-i

b. *Tomba-na Ibemma-bu phajaba thAja-i

PROPERTIES OF CGs...

iv). CGs can never occur as complements of subject raising verbs although they can occur as a single constituent in the subject position of raising predicate:

English:

(20) a. *John appears [liking Mary]

b. [(John) talking to Mary] seems impossible.

Manipuri:

(21) a. *TamchA-na pAmba Chaobi-bu mAlli
Tomcha-ERG liking Chaobi-ACC appear

b. TomchA-na ChAobi-bu pAmba mAlli
TomchA-ERG ChAobi-ACC liking appear
It appears [that Tomcha likes Chaobi]

v). The subject position og a CG must be filled in the course of derivation, either by a lexical DP (a) below, or by a pure expletive (b) below to satisfy the EPP requirement:

(22) a. Paul prefers [~~Paul~~ swimming in the morning].

b. Bill enjoys [there being many people at the party]

Manipuri:

(23) Khomei [ayuk-ta ~~Khomei~~ iroiba] pamja-I
Khomei morning-LOC swimming prefer

CLAUSAL GERUND

4.2 Deriving CG:

Pires (2006: 39) proposed three hypothesis regarding the properties of CGs:

(24). a. The inflectional head corresponding to -ing in English (-pa/-ba in Manipuri) in CGs carries a feature specification that forces the occurrence of CGs in positions accessible to Case valuation;

b. In the derivation of a CG, the Case feature of its external argument DP can be valued within the CG itself (25a & 26a below);

c. The external argument DP can move out of the CG before the CG can value the Case feature of this DP. This yields a null-subject CG (a CG with a control PRO subject, in standard term) (25b & 26b below).

(25). a. Sue prefers [John/him swimming]

b. John prefers [swimming]

(26). a. Tomba-na [Khomei-bu/mA-bu irujaba] pAmmi

b. Tomba-na [irujaba] Pammi

Under this approach, the head T of the CG itself will be a goal for Case valuation, i.e., the -ing in English and the suffix -pa/-ba in Manipuri.

DERIVING CGs…

NOTE: Here, the adopted approach to overt syntax explores certain core aspects of the architecture proposed in Chomsky 2000, 2001 in terms of phrase structure, Case, Φ-feature and A-movement to subject position. Case and Φ-feature valuation are taken to apply as a consequence of the operation *Agee:*

Agree "establishes a relation (agreement, case checking) between an LI [lexical item] α and a Feature F in a search space (its [the LI's] domain)"

(Chomsky 2000:102) ; (LI α is the Probe; Feature F is the Goal).

Match: Probe and Goal need to have a subset of their features in common (Φ-feature here).

Now let us derive the following CG:

(27a) John prefers [~~John~~ swimming] (English)

Tomba [~~Tomba~~ iroiba] pAmmi (Manipuri)

It is proposed that the null subject in such cases results from A-movement of the embedded CG subject to the matrix clause. The Θ-roles can be assigned through movement and not only by first merge (cf. Boskovic 1994, Lasnik 1995, Boskovic and Takahashi 1998). Θ-roles can also be assigned in the course of derivation, and are satisfied not in a configuration, but in a set of configurations (i.e. transformationally).

DERIVING CGs…

(27) Tomba iroiba pAmmi

a. $[_{T'}$ AGR $[_{vP}$ Tomba iroiba]]
[Φ/$Case_{AGR}$ Θ/Case

b. $[_{TP1}$ Tomba $[_{T'}$ AGR $[_{vP}$ ~~Tomba~~ iroiba]]]
[~~EPP/ Φ~~/$Case_{AGR}$ [Θ/Case]]]

c. $[_{vP}$ Tomba $[_{v'}$ pAmmi $[_{TP1}$ ~~Tomba~~ $[_{T'}$ AGR $[_{vP}$ ~~Tomba~~ iroiba]….
[2Θ/Case [~~Case~~ $_{AGR}$ [~~EPP~~/ ~~Φ~~ [Θ]….

d. $[_{TP2}$ Tomba $[_{T'}$ $[_{vP}$ ~~Tomba~~ $[_{v'}$ pAmmi $[_{TP1}$ $[_{T'}$ AGR $[_{vP}$ iroiba]….
[~~Φ~~/~~Case/~~ ~~EPP~~ [2Θ [~~Case~~$_{AGR}$ [~~EPP~~ /~~Φ~~ [vP]….

As *Tomba* enters Match/Agree with AGR in (27b), *Tomba* values the Φ-set of AGR by Agree and moves to Spec TP1 for EPP satisfaction. But, since AGR still has an uninterpretable Case feature at the point in (27b), Case valuation of the embedded subject DP cannot yet take place. When the matrix v is inserted in the derivation, the embedded CG is assigned the propositional internal Θ-role of the matrix verb (27c). When the matrix v enters the derivation, it attracts the embedded DP *Tomba* and assigns an experiencer Θ-role to it. The matrix v then Matches/Agrees in Φ-features with the AGR in CG and values the uninterpretable Case feature (C_{AGR}) that AGR still carries (27c). Finally, *Tomba* moves from matrix [Spec, vP] to check/value its own uninterpretable Case feature and the EPP and Φ-features on TP2 (27d).

5.0 FEATURE SPECIFICATION FOR FURTHER LEVEL OF TAGGING

✓ We now realize that a verbal noun can project at least three strctures:

1. **Derived nominal structure,** having the configuration of:

DP-GEN VN > ((khungang-gi/NP.gen mAngaba/VN))/DN

2. **Gerund structure,** having the configuration of:

DP-GEN VN-ACC > ((khungang-bu/NP.acc mAngnaba-bu/VN.acc))/GND

3. **Clausal gerund structure**, having two configurations of:

a. PRO VN > mAngnaba/CG

b. DP VN > ((Tomba /NPP cAba/VN))/CG

For the purpose of disambiguity, we can merge Gerund and Clausal into single category as **Gerund** only.

5.0 CONCLUSION

In conclusion, if only parts of speech category is supposed to be considered in POS tagging without introducing other higher level categorical names such as Noun Phrase, Verb Phrase, Adjective Phrase, adverbial Phrase, Prepositional Phrase, Gerund, Derived Nominal, TP, CP etc., it is better to suggest that only Verbal Noun ought to be introduced; and the so-called Verbal Noun should be placed under Non-finite verbs.

ACKNOWLEDGEMENT

TO DR. RAMAMOORTHY

HEAD, LDC-IL, CIIL, MYSORE.

Dr. Richa, Attreyee, Arundhati, Thennarasu, Shahid, Dr. Jitendra, Yoonus, Vadieval, Rajesh, Manasa, Bharat, Sarita, Poonam, Anand, Premila, Bidyarani, Sonali, Vijalaxmi, Mansoor, Farson, Bridul, Umesh, Rupesh, Sarbjit, Arunkumar, Dinesh, Atul, Kulwant, Pramod, Khirod, Satyeindra, Amrudha, Ashok, Deepti, Geeta, Sign Rajesh, …..

My memory fails…..

To you all of LDC-IL.

THANKS