Nominalism and Lexical Categories in Austronesian: A case study in Evolutionary Syntax

This talk in a nutshell:
I. Intro: on the Noun-Verb distinction  
II. Semantic coercion versus categorial conversion in Tagalog and beyond  
III. The nominalist hypothesis in Austronesian  
IV. The synchronic payoff  
V. The redevelopment of Verbs in Indonesian languages  
VI. Conclusion: prospects for Evolutionary Syntax

I. Introduction: On Nouns and Verbs

- Identifying Nouns and Verbs in Tagalog and other Philippine languages is notoriously thorny. Syntax doesn’t offer many clues, notionally verbal morphology applies quite freely to notional nouns (without extra derivation) and notional verbs have certain nominal properties.

- Here, we will argue that there really is no good evidence for positing a separate category of verbs in Standard Tagalog, (Gil 1993, 2000; Himmelmann 1991, 1993, 2007), and this extends to many other Philippine languages.

- We’ll refer to intuitive/notional categories as “Verbs” and “Nouns”. But the real linguistic categories of interest are:

  ![Diagram of linguistic categories]

  - We will only require the most rudimentary **SemCat** distinctions here: Event versus Entity. These can be defined formally in predicate calculus terms as in (2) or on the basis of Langacker-type schemas as in Fig. 1.

    John loves Mary

    $\exists e (\text{Love}(e) \& \text{Lover}(e, john) \& \text{Loved-one}(e, mary))$

    $e = \text{Davidsonian event variable}$
In the schema of <enter> there is motion and a timeline. The definition of an Event is a lexeme which includes a timeline. An Entity can be conceived of as having no timeline or constituting a slice of a timeline.

- **MorphCat** encompasses morphological marking of any nature. While Aspect is a SemCat, the surface marking of [ASPECT] is a morphological feature. So the bare root <enter> has the lexical structure of an Event but is not [+ASPECT].

- **PhraseCat** includes the marking and positional properties associated with particular syntactic constituents. For our purposes, the position and marking of predicates and arguments will be the PhraseCats of primary interest.

- In languages with a high level of categorization, the three categories typically overlap like this:
  
  - *Entity (SemCat)*
  - *Number, Definiteness (MorphCat)*
  - *Argument position (PhraseCat)*
  - *Event (SemCat)*
  - *Aspect, Tense (MorphCat)*
  - *Predicate position (PhraseCat)*

  This overlap has led to much confusion and conflation in the literature. Complicates the analysis of mismatches (cf. Hopper & Thompson 1984, 1985; Hengeveld 1992:58; Croft 1991, 2000, 2001; Gil 2000). In particular, SemCat, MorphCat and PhraseCat are all commonly taken as evidence for lexical categories without rigorous argument.

  Root level and word level acategoriality must be kept distinct as one does not entail the other (Lehmann 2006).

- **Root categorialization**: Restricted morphological potentials for roots, e.g., in a language where Event denoting roots cannot take Number morphology and Entity denoting roots cannot take Aspect morphology.

<table>
<thead>
<tr>
<th>Restricted √s:</th>
<th>Unrestricted √s:</th>
</tr>
</thead>
<tbody>
<tr>
<td>√Event-ASP</td>
<td>√Event-ASP</td>
</tr>
<tr>
<td>∗√Entity-ASP</td>
<td>√Entity-ASP</td>
</tr>
<tr>
<td>∗√Event-NUM</td>
<td>√Entity-NUM</td>
</tr>
<tr>
<td>√Entity-NUM</td>
<td>√Entity-NUM</td>
</tr>
</tbody>
</table>
Word categorialization: Restricted syntactic potential for words, e.g., in a language where Aspect marked words cannot appear in argument position and Number marked words cannot appear as bare predicates.

<table>
<thead>
<tr>
<th>Restricted Wds:</th>
<th>Unrestricted Wds:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred[Word·ASP]</td>
<td>Pred[Word·ASP]</td>
</tr>
<tr>
<td>*Subj[Word·ASP]</td>
<td>Subj[Word·ASP]</td>
</tr>
<tr>
<td>Pred[Word·NUM]</td>
<td>Pred[Word·NUM]</td>
</tr>
<tr>
<td>Subj[Word·NUM]</td>
<td>Subj[Word·NUM]</td>
</tr>
</tbody>
</table>

II. Noun and Verb in Tagalog

Tagalog background: Predicate initial. All arguments are case marked (T1). Event denoting predicates are obligatorily marked with Voice and Aspect. The voice system (T2) is complex, being able to select a large number of argument and adjunct type constituents as subject. The constituent selected by the voice morphology is put in the nominative case and other (non-directional) arguments are put in the genitive case. Directional arguments not selected by the voice morphology are put in the oblique case.

Table 1. Tagalog case markers

<table>
<thead>
<tr>
<th>PERSONAL NAME</th>
<th>GENERAL</th>
<th>PERSONAL NAME</th>
<th>GENERAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>PL</td>
<td>SG</td>
</tr>
<tr>
<td>NOMINATIVE</td>
<td>aŋ</td>
<td>sì</td>
<td>sina</td>
</tr>
<tr>
<td>GENITIVE</td>
<td>naŋ</td>
<td>ni</td>
<td>nina</td>
</tr>
<tr>
<td>OBLIQUE</td>
<td>sa</td>
<td>kay</td>
<td>kina</td>
</tr>
</tbody>
</table>

Table 2. Tagalog voice system (súlat 'write')

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Voice</th>
<th>ACTOR</th>
<th>PATIENT</th>
<th>LOCATIVE</th>
<th>CONVEYANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASIC</td>
<td>&lt;um&gt;</td>
<td>sumúlat</td>
<td>sulátin</td>
<td>isúlat</td>
<td></td>
</tr>
<tr>
<td>PERFECTIVE</td>
<td></td>
<td>magsulat</td>
<td>sulátan</td>
<td>sinúlat</td>
<td>isinúlat</td>
</tr>
<tr>
<td>IMPERFECTIVE</td>
<td></td>
<td>sumúsulat</td>
<td>sinúsulátan</td>
<td>sinúsulátan</td>
<td>isinúsulát</td>
</tr>
<tr>
<td>PROSPECTIVE</td>
<td></td>
<td>súsulat</td>
<td>súsulátin</td>
<td>súsulátan</td>
<td>isúsulát</td>
</tr>
</tbody>
</table>

Voice system in plain declaratives

(1)a. s<um>úlat=ka naŋ=líham b. sulát-an=mo anŋ=dindin

Write a letter!’ ‘Write on the wall!’

c. sulát-in=mo anŋ=líham d. i-panúlat=mo anŋ=lápís

Write the letter!’ ‘Write with the pencil!’

e. i-súlat=mo naŋ=tulà anŋ=asáwa=mo

Write a poem for your spouse!’
The voice system in interrogatives

(2)a. Ano aŋ=s<in>úlat-∅=mo?
   what NOM=<RL>write-PV=2S GEN
   ‘What did you write?’

b. Sino aŋ=s<um>úlat nito?
   who.NOM NOM=<AV RL>write this GEN
   ‘Who wrote this?’

c. Sino aŋ=i-s<in>úlat=mo naŋ=tulà?
   who.NOM NOM=CV<RL>write=2S GEN GEN=poem
   ‘Who did you write a poem for?’

Two flavors of acategoriality in Tagalog:

Unrestricted words (aka omnipredicativity)

(3)a. [nag-tá~trabáho] aŋ=[laláki]
   AV-INCM-work NOM=man man NOM=AV-INCM-work
   ‘The man is working’
   ‘It’s a man who’s working’

b. [laláki] aŋ=[nag-tá~trabáho]
   AV=INCM~work NOM=AV=INCM~work
   ‘A man who’s working’

Unrestricted roots

   woman=LNK AV-INCM-work AV-INCM-work=LNK woman
   ‘working woman’
   ‘working woman’

   <AV>INCM~road NOM=youth <AV>INCM~youth
   NOM=road
   ‘The youth passes by.’
   ‘The road gets younger.’

But acategoriality ≠ anarchy. There are still bad combinations:

(6)a. mag-bigay b. i-bigay c. bigy-an d. *bigay-in
   AV-give CV-give give-LV give-PV
   ‘x to give’ ‘to give x’ ‘to give to x’

Patient voice implies undergoer affectedness. Because giving does not affect the
theme, the –in affix is semantically incongruous. Cf. Ballard (1974) on Ibaloy
(Northern Philippine):

“The meaning of a verb is the combination of the meaning of the verb root plus the
meaning added by the affixation. ...A root will occur with those affixes whose
meaning is compatible with its own meaning.” (via Huang & Huang 2007:425 fn.3)
De Guzman argues on the basis of other morphological operations that there exists a N-A-V distinction in Tagalog:

“...the specific forms – simple roots, typical affixes – the type of reduplication, and/or stress modification that apply to each word class corresponding to specifiable meanings within the class are characteristic to each category. Thus, every lexical derivation rule has to specify the category of its input and the category of its output, even when there is no change in category.” (DeGuzman 1995:312)

 Iterative, Intensive and Moderate morphology as evidence for categoriality:

(7) a. káin naŋ=káin b. táwa naŋ=táwa
eat GEN=eat laugh GEN=laugh
‘constantly eating’ ‘constantly laughing’

(8) a. *báhay naŋ=báhay b. *ma-ganda naŋ=ma-ganda	house GEN=house ADJ-beauty GEN=ADJ-beauty

 Iterative reduplication is an operation on the lexeme’s aspectual structure/timeline, as illustrated by the following Langackerian representation (Langacker 1999, I-wen Su & Huang 2006):

If the lexeme in question has no timeline, iterativity will be uninterpretable. In fact, “Noun” roots with directional semantics are possible here. And “Noun” roots such as <báhay> can take iterative reduplication once Voice+Aspect are added (as also noted by De Guzman).

(9) a. daan naŋ=daan b. nag-báhay naŋ=nag-báhay
road GEN=road AV.RL-house GEN=AV.RL-house
‘passing by’ ‘constantly made/settled in houses’

c. nag-má~ma-ganda naŋ=nag-má~maganda
AV.RL-INC~ADJ-beauty GEN=AV.RL-INC~ADJ-beauty
‘constantly pretending to be beautiful’

This suggests the following hypothesis: Morphosyntactic operations may make direct reference to semantic features and these features can either be present at the √-level or be coerced at higher levels. Aspectuality is present on Event type and directional roots and is coerced by aspectual morphology.

Coercion is a widespread phenomenon in Tagalog and beyond:

 Coercion of Entity > Property via adjectival ma-
<ganda> ‘beauty’ Property; <pérà> ‘money’ Entity
We can further note that the lack of the Noun-Verb distinction is in part due to the nouniness of the Voice-Aspect inflected words.
(20) a. ma-šuŋkot na b<um>alik
   ADJ-sad   LNK <AV.RL> return
   ‘returned sadly’ or ‘sad returned one’

   b. ma-šuŋkot na bátà
   ADJ-sad   LNK child
   ‘sad child’

(21) a. Mayroon / walà / ma-rámi=ŋ sulat
   EXT / NEG.EXT / ADJ-many=LNK letter
   ‘There are / there are no / there are many letters.’

   b. Mayroon / walà / ma-rámi=ŋ s<in><úlat-∅
   EXT / NEG.EXT / ADJ-many=LNK <RL>letter-PV
   ‘There are / there are no / there are many things written.’

   c. Mayroon / walà / ma-rámi=ŋ nag-sú~sulat
   EXT / NEG.EXT / ADJ-many=LNK <AV.RL>INCM~letter
   ‘There are / there are no / there are many who write.’

III. The nominalist hypothesis (Starosta, Pawley & Reid 1982)

- The voice forms which are now more commonly found as matrix predicates
developed from nominalizations. These may have been required in embedded and
subordinate contexts.

- The original Austronesian verbs were the progenitors of what are termed
“dependent forms” in certain modern languages.

Table 3. Independent and dependent affixes (Ross 2002)

<table>
<thead>
<tr>
<th></th>
<th>ACTOR</th>
<th>PATIENT</th>
<th>LOCATIVE</th>
<th>CONVEYANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDICATIVE</td>
<td>&lt;um&gt;</td>
<td>‘in’</td>
<td>‘an’</td>
<td>‘i’</td>
</tr>
<tr>
<td>DEPENDENT</td>
<td>Ø</td>
<td>‘a’</td>
<td>‘i’</td>
<td>‘an’</td>
</tr>
</tbody>
</table>

- Although not present in Standard Tagalog, the dependent paradigm survives in
Southern varieties of Tagalog (e.g. Batangas) in addition to a large number of
other languages throughout the Philippines.

- Morphosyntactic differences between dependent and indicative forms can be
explained as Verb vs. Noun. In Batangas Tagalog and other Philippine
languages, imperative addressee of dependent forms are obligatorily absent:

(22) Buks-i(*=mo) an= bentána!
    open=LV.DEP=2S.GEN NOM=window
    ‘Open the window!’

- Unlike indicative imperatives, dependent forms cannot be embedded in a DP,
quantified or modified directly by adjectives:

(23) Bentána=na=laŋ an=buks-an=mo!
    window=CMP=only NOM=open-LV=2S.GEN
    ‘Open the window!’ (Lit. ‘Window is your one to open!’)
These are 3 good distributional arguments for treating the indicatives as nominal and the dependents as verbal.

ERG = GEN arises from the agent marking in nominalized subordinate clauses which was then extended to matrix clauses. Because agents of nominalizations are generally introduced as possessors, this reanalysis brings about an agent-possessor (i.e. ergative-genitive) syncretism.

IV. Diachronic solutions to Synchronic problems

The intuition that verbs have strong nominal characteristics has been recognized by historical linguists but has not made it into a single modern syntactic analysis of any AN language.

The work of Alana John’s (1992) on Inuktitut (Canadian Eskimo) offers a cogent formal exposition of the Nominalist Hypothesis and can be applied to AN. The basic elements of the analysis are the following:

i. VERB roots are unable to project a VP
ii. Transitive clauses require passivization via a passive participle
iii. The passive morpheme is a nominalizer

a. kapi-jaq
   stab-PASS.PART
   ‘the stabbed one’

b. anguti-up kapi-ja-a-∅
   man-GEN stab-PASS.PART-3S/3S
   ‘The man’s stabbed one.’

c. anguti-up nanuq kapi-ja-a-∅
   man-GEN bear(ABS) stab-PASS.PART-3S/3S
   ‘The man stabbed the bear.’

d. angut ani-juq
   man(ABS) go.out-INTR.PART.3s
   ‘The man went out’
Under this analysis, it appears that the genitive agent attaches within the domain of modification, while the absolutive attaches in the domain of predication.

In Tagalog resultative nominals, genitive phrases are also interpreted as agents and the subject is part of an equational structure with the initial bare predicate (cf. Bloomfield 1917)

Kítà ŋ laláki ŋ baŋka
visible GEN=man NOM=boat 'The man sees/saw the boat'
(Lit. 'The boat is the man’s visible one')

The crucial aspect of the above structure is that the dependents of the predicate are contained within a single domain of modification.

**Prediction**: Genitive agents should have the hallmarks of modifiers

In the preposed pronominal construction the arguments are attached to the following constituent by the linker. The linker is a diagnostic for modification in Philippine languages, demarcating the edges of DP-internal constituents.

(29) \[ \text{DP[Itō=} \text{ŋ ma-laki=} \text{ŋ áso=} \text{ŋ ito]} \]
NOM.this=LNK ADJ-big=LNK dog=LNK NOM.this
'This big dog'

Preposed agents are attached with the linker, while this is impossible for the nominative subject.

(30) \[ \text{DP[Ákin=} \text{ŋ na-kítà]} \]
1S.OBL=LNK PV.STA.see NOM=woman
'I saw the woman.' (cf. Naylor 1980, 1995)
This also captures the reason why Austronesian languages have stringent restrictions on extraction.

- The famous restriction on extraction in (conservative) AN langs has been dubbed “subjects-only” but this is a misnomer. The restriction is demonstrated below with topicalization, which does not require altering the predicate-argument structure of the sentence:

(32) a. B<in>ili naŋ=babáe aŋ=libro kahápon <PV.RL>buy GEN=woman NOM=book yesterday
   ‘The woman bought a book yesterday.’

b. [Ang=libro], ay b<in>ili- Ø naŋ=babae aŋ=libro tì yesterday TOP <RL>buy-PV GEN=woman NOM=book

c. [Kahápon], ay b<in>ili- Ø naŋ=babae aŋ=libro tì yesterday TOP <RL>buy-PV GEN=woman NOM=book

d. *[Naŋ=babáel], ay b<in>ili-Ø tì aŋ=libro kahápon GEN=woman TOP <RL>buy-PV NOM=book yesterday

- **If the genitive agent is a DP-internal possessor, we expect it to be restricted!** Cross-linguistically, possessors are highly constrained in their extraction possibilities, (cf. Ross 1967; Keenan & Comrie 1977, 1979a, 1979b; Gavruseva 2002 *inter alia*).

  Accessibility Hierarchy (Keenan & Comrie 1977)
  
  SU > DO > IO > OBL > GEN > OCOMP

  This solves a paradox encountered by the AH. Keenan & Comrie were forced to assimilate the absolutive argument to the subject in the (mostly Austronesian) ergative languages in their sample.

- **Some possessor extraction data**

(33) a. [et=ha=byt šel mi] raita tì? Hebrew
   OBJ=DEF=house of who see.PST.2S
   ‘Whose house did you see?’

b. *[(šel) mi raita], [et=beyt-o tì]? (pied piping of DP of who saw.PST.2S OBJ=house-3S.GEN obligatory)

c. *[šel mi raita], [(et=ha=)byt tì]? of who saw.PST.2S OBJ=DEF=house
(34) a. [beːt miːn] šuft t?  
    house who see.PST.2s  
    ‘Whose house did you see?’

    b. *miːn šuft [beːt t]?  
    who see.PST.2s house  
    (pied piping of DP obligatory)

(35) [Whose pictures] did you see t?  
    *[Whose], did you see [t, pictures]?  
    *[Of whom], did you see pictures t?  
    (but, [Whom] did you see pictures [of t] ? (conditioned))

    What did John write about Nixon? He wrote it (=a book) about Nixon.  
    *What did John see of Nixon?  *He saw it (=a picture) of Nixon.

    [VP write [DP a book ] [VP about Nixon ]]  
    [VP see [DP a picture [VP of Nixon ]]]  
    (Chomsky 1977)

(36) a. (a) Mari-Ø vendég-e-Ø  
    the Mari-N/G guest-POSS.3s  
    ‘Mary’s guest’

    b. *Mari-Ø a vendég-e-Ø  
    Mary-N/G the guest-POSS.3s

    c. Mari- nak a vendég-e-Ø  
    Mary-DAT the guest-POSS.3s  
    (GEN>DAT case shift w/extraction)  
    ‘Mary’s guest’  
    (Szabolcsi 1983/1984)

(37) a. [Gambar=nya siapa] kamu=lihat t?  
    picture=3s.GEN who 2=see  
    ‘Whose picture did you see?’

    b. *Siapa kamu=lihat gambar=nya?  
    (Pied piping of who 2=see picture=3s.GEN entire DP obligatory)

(38) a. baloy ru ruandu’  
    house GEN woman  
    ‘a woman’s house’ (Prentice 1971:180)

    b. baloy=min  
    house=2s.GEN  
    ‘your house’ (Prentice 1971:181)

    c. ruandu’ ra baloy  
    woman LNK house  
    (GEN>NOM case shift w/symmetrical linking)  
    ‘a WOMAN’S house’ (Prentice 1971:205)
d. akaw ra baloy
   2s LNK house
   'YOUR house' (Prentice 1971:205)

† In Tagalog, several types of arguments are marked by naŋ GEN, not all of them possessors (cf. Kroeger 1993:40-7). Here, restricted extraction is a result of position and function, not case. We expect genitive case not to effect the extractability of elements in the predicational domain:

(39) a. <Um>alis=sila naŋ=ala-úna
   <AV.RL>leave=3P.NOM GEN=o’clock-one
   ‘They left at one o’clock’

b. Naŋ=ala-úna ay <um>alis=sila
   GEN=o’clock-one TOP <AV.RL>leave=3P.NOM
   ‘At one o’clock, they left’

(40) a. Walà=ako=ŋ g<in>awà-Ø nito=ŋ maña=huli=ŋ áraw
   NEG.EXT=1S.NOM=LNK <RL>do-PV GEN,this=LNK PL=last=LNK day
   ‘I haven’t done anything, these last (few) days.’

b. Nito=ŋ maña=huli=ŋ áraw ay walà=ako=ŋ g<in>awà-Ø
   GEN,this=LNK PL=last=LNK day TOP NEG.EXT=1S.NOM=LNK <RL>do-PV
   ‘These last (few) days, I haven’t done anything.’

(41) a. Na-húlog=siya naŋ=hindi=niya na·maláy=an
   PV.STA.RL·fall=3S.NOM GEN=NEG=3S.GEN STA.RL·conscious-LV
   ‘He fell without noticing it’

b. Naŋ=hindi=niya na·maláy=an ay na-húlog=siya
   GEN=NEG=3S.GEN STA.RL·conscious-LV TOP PV.STA.RL·fall=3S.NOM
   ‘Without noticing it, he fell.’

Interestingly, a distinction exists in GEN marked adverbials. Among temporal adverbs, punctuals are extractable, while duratives are not:

(42) a. Nag-áral=sila naŋ=isa=ŋ óras
   AV.RL·study=3P.NOM GEN=one=LNK hour
   ‘They studied for an hour’

b. *Naŋ=isa=ŋ óras ay mag-áral=sila
   GEN=one=LNK hour TOP AV.RL·study=3P.NOM

Also, frequency adverbs are extractable while manner adverbs are not:

(43) a. T<um>akbo=sila naŋ=ma-dalas
   <AV.RL>run=3P.NOM GEN=ADJ·often
   ‘They run often.’
b. **Ma-dalas** ay t<um>akbo=sila
   ADJ-often TOP <AV.RL>run=3P.NOM
   ‘Often, they run’

(44) a. T<um>akbo=sila **naŋ=ma-bilis**
   <AV.RL>run=3P.NOM GEN=ADJ-speed
   ‘They ran quickly.’

b. *Ma-bilis** ay t<um>akbo=sila
   GEN=ADJ-speed TOP <AV.RL>run=3P.NOM

If topicalization of phrasal adverbs is generally permitted by the grammar, we can derive the facts based on the nominalist hypothesis. Only punctual adverbs relate to the entire predicational domain; durative adverbs only relate to the smaller, modificational domain. Extraction of duratives thus violates the same island constraint seen above with possessors.

V. The disintegration of nominalism in Indonesia

❖ Outside the Philippines, the nominal system described here breaks down. Further support for the nominalist hypothesis comes from the fact that throughout various Indonesian subgroups, the inherited features discussed above are lost and many of the same innovations arise.

❖ The linker, the primary demarcator (and acquisition cue) of the modification domains, disappears. This results in the creation of a real (i.e. category particular) **relative marker**.

(45) a. Anjiŋ besar  
    dog big

b. Anjiŋ **yaŋ** besar  
   dog RELT big
   ‘a big dog’

   b. Anjiŋ **yaŋ** besar  
   dog RELT big
   ‘a dog which is big’

c. Anjiŋ *(yaŋ)* ku=lihat
   dog RELT 1s=see
   ‘the dog I see’

❖ Case markers, the primary indicators of which phrases are in which domain, erode. Nominative pronouns can typically serve as AV patients unlike in Philippine languages where pronominal AV patients must be marked with oblique case. Case no longer signals the subject of the predicational domain.

(46) a. Aku me-lihat **kamu**
    1s(NOM) AV-see 2s(NOM)
    ‘I see you’

(47) **ŋ-ita** kaw aku...
    AV-see 2s(NOM) 1s(NOM)
    ‘I see you’  (Donohue & Brown 1999:71)
The agents of patient voice verbs are no longer obligatorily genitive modifiers. They can be introduced as obliques, signalling the emergence of true passives.

(48) Ini buku kamu
this book 2(NOM)
‘This is your book’

(49) Ni-kokko’=a’ ri meoŋ=ku
PASS-bite=1S.NOM PREP cat=1S GEN
‘I was bitten by my cat’ (Jukes 2006:254)

(50) Mbe’e ede ra’nduku ba ompu sia
goat that PASS.REAL-hit by grandfather 3sg
‘The goat was hit by his/her grandfather’ (Arka 2002)

(51) Tu’ da-kerja ulih dua iku’ nsia
this PASS-work by two CLASS human
‘This is done (later) by two persons.’ (Tjia 2007:152)

(52) Ami ongga le hia
1p.ex hit by 3s
‘We were hit by him/her’ (Arka & Kosmas 2005)

(53) Lôn ka geu-côm lé-gopnyan
1p IN 3-kiss OBL she
‘I was kissed by her.’ (Durie 1988, Lawler 1988, Asyik 1987)

Voice inflected elements can now license applicatives, person agreement, ∅ imperative addressees, i.e., they are real Vs.

(54) Ini yang ku=tulis
this RELT 1s=write
‘This is what I wrote’

(55) Aku men-ulis-kan kamu sajak
1s AV-write-APPL 2 poem
‘I write a poem for you’

(56) Bib n=pun-ak kolay peda
Bib 3s=kill-APPL snake machete
‘Bib killed the snake with a machete’ (Bowden 2001:122)

(57) tau ku=buntul-i=a...
person 1S.ERG=meet-APPL=DEF
‘the person that I met’ (Jukes 2006:239)
(58)  Ia meli-\textit{an} Nyoman umah  \textit{Balinese}
3 AV.buy-\textit{APPL} name house
‘(S)he bought a house for Nyoman’

(59)  Masak sayur=nya! \textit{Indonesian}
cook vegetable=3S.GEN
‘Cook the vegetables!’

(60)  Keo=\textit{a}! \textit{Selayarese}
call=1S.NOM
‘Call me!’

- Person agreement, which in most cases develops from the genitive set pronouns, deviates from possessor marking, as in many languages of Sulawesi (Noorduyn 1991:148-9):

\begin{align*}
\text{na-}\text{hilo}=a & \quad \text{tomi}=\text{ku} & \text{Uma} \\
3S.\text{ERG-see}=1S.\text{GEN} & & \\
\text{na-}\text{cini}=\text{ka}' & \quad \text{ballak}=\text{ku} & \text{Makassarese} \\
\text{na-}\text{kita}=\text{na}' & \quad \text{banua}=\text{ngku} & \text{Sa’dan} \\
\text{la-}\text{loja}=\text{aku} & \quad \text{sapo}=\text{ku} & \text{Barang-barang} \\
\text{a-}\text{komata}=\text{aku} & \quad \text{banua}=\text{ku} & \text{Wolio} \\
\text{no-}\text{toa}=\text{aku} & \quad \text{laika}=\text{ngu} & \text{Tolaki} \\
\end{align*}

- Occasionally copulas and indefinite pronouns crop up as well:

(61)  Ini \textit{adalah} guru saya \textit{Indonesian}
this COP teacher 1s
‘This is my teacher’

(62)  Aku men-erja-kan \textit{sesuatu} \textit{Indonesian}
1s AV-work-\textit{APPL} something
‘I’m doing something’

(63)  %Ada di-kerja-kan \textit{Indonesian}
EXT PASS-work-\textit{APPL}
(For, ‘He worked on something’)

- Topicalization of actor voice patients is permitted, as the verb now “governs” its object:

(64)  Pa\textit{n}uman itu, gaukan no-gutu \textit{Totoli}
story DIST king AV.RLS-make
‘This announcement, the king made’ \textit{(Himmelmann 2006:142)}
V. Conclusion – An evolutionary approach

- A new taxonomy for ergative languages (cf. Palancar 2002):

  'With the ergative, type 2 syncretism in our sample most typically joins it with the genitive, as in the Tacanan language Araona, the isolate Burushaski, Lak and the Tibeto-Burman language Limbu. It is likely that this is not a random choice, in that there are languages which have cases which inherently combine the functions of ergative and genitive (e.g. the relative case of the Eskimoan languages). Such constructions may have their origin in nominalizations, with the agent expressed by the genitive. However, although diachronic explanations may be found, it is unlikely that a direct, synchronic motivation can be demonstrated for most type 2 patterns' (Baermann et al 2005:52, emphasis mine)

- Restrictions on extraction in ergative type languages appear to closely mirror the restrictions in their source constructions. As a result, extraction in ergative-genitive languages is much more restricted than in ergative-instrumental languages (cf. Manning 1996).
(65) Gizona-k neska-rí pelota-∅ eman d-ío-∅  
man-ERG girl-DAT ball-ABS given 3S.ABS-3S.DAT-3S.ERG  
‘The man has given the ball to the girl.’ (Breitschneider 1979:374)

Relativization of Abs: [Gizona-k neska-rí eman d-ío-∅ n]  
man-ERG girl-DAT given 3S.ABS-3S.DAT-3S.ERG ball  
‘The ball that was given by the man to the girl.’

Relativization of Dat: [Gizona-k pelota-∅ eman d-ío-∅ n]  
neska  
man-ERG ball-ABS given 3S.ABS-3S.DAT-3S.ERG girl  
‘The girl that was given the ball by the man.’

Relativization of Erg: [Neska-rí pelota-∅ eman d-ío-∅ n]  
gizona  
girl-DAT ball-ABS given 3S.ABS-3S.DAT-3S.ERG man  
‘The man who gave the ball to the girl.’

We also predict that Indonesian languages which have redeveloped verbs and possess non-genitive agents should behave like *instrumental* ergative languages. This is borne out by several languages, e.g.:

(66) [Ulih dua iku’ nsia] tu’ da-kerja  
by two CLASS human this PASS-work  
‘This is done (later) by two persons.’ (Tjia 2007:152)

In the majority of Indonesian languages (excluding CEMP) the typical extraction restrictions of the Philippine system exist with clefting, but not with topicalization. Also,

‘...non-subject undergoer arguments in actor voice constructions...can usually be topicalized without any problems but non-subject actor arguments in undergoer voice constructions.” (Himmelmann 2006:143)

The way in which nominalism offers an explanation for a constellation of typological facts in Philippine languages suggests that it is crucial to look to historical facts for explanation.

More importantly, the way in which certain consequences of nominalism survive in Indonesian languages long after nominalism has been eliminated suggests an evolutionary approach to syntax along the lines of Blevins (2004)

Evolutionary Phonology:

Principled diachronic explanations for sound patterns have priority over competing synchronic explanations unless independent evidence demonstrates, beyond reasonable doubt, that a synchronic account is warranted. (Blevins 2004:237)
In fact, we still have very little idea what the threshold for irregularity is in first language acquisition. Ever since the shift towards a synchronic view (Sausurre 1916), it has often been assumed that linguistic systems must ‘make sense’ on some synchronic level.

But if the evidence for “hyperlearning” and Poverty of the Stimulus is weaker than previously thought (Pullum & Scholz 2002, Scholz & Pullum 2002 inter alia) systematicity could be enforced much less stringently.

SOURCES
Blevins, Juliette. 2004. Evolutionary Phonology. CUP.
Donohue, C, Donohue, M. forthcoming. The case of possessors and subjects. ms.
Gil, David. 2000. Syntactic categories, cross-linguistic variation and universal grammar.
in Approaches to the Typology of Word Classes. Petra M. Vogel and Bernard
Comrie (eds.). Mouton de Gruyter.
Himmelmann, Nikolaus. 2006. Typological Characteristics. In A. Adelaar and N.
Himmelmann (eds.), The Austronesian Languages of Asia and Madagascar.
London: Curzon.
Tagalog. Munchen: Fink
Himmelmann, N. 1996. Person marking and grammatical relations in Sulawesi. H.
Steinhauer (ed.) Papers in Austronesian Linguistics No. 3. Canberra: Pacific
Linguistics.
Himmelmann, N. 2007. Lexical categories and voice in Tagalog, in Voice and
Grammatical Functions in Austronesian Languages. Stanford: CSLI
Linguistic Inquiry 8.63-69.
Keenan, E. & Comrie, B. 1979a, Noun Phrase Accessibility Revisited. Language, 55:3,
pp. 649-664.
Language 55. pp.333-51
the third LFG conference. Brisbane.
Lehmann, Christian. Roots, stems and word classes. Paper given at ‘Universality and
Stanford: CSLI
Naylor, P.B. 1995. Subject, Topic and Tagalog Syntax, in Subject, Voice and Ergativity,
161-201, London: SOAS.
Austronesian Studies, Papers from the 2. Eastern Conference on Austronesian
“Subject, Topic, and Tagalog syntax”. In: Benett, David, Bynon, Theodora and
Noorduyn, J. 1991. The languages of Sulawesi. in Papers in Austronesian Linguistics
Oda, Kenji. 2005. ‘V1 and wh-questions: A typology.’ In Verb first, (ed.) Andrew Carnie,
Akademie Verlag


