The typology of nominal licensing in Austronesian voice system languages

Michael Yoshitaka Erlewine, Theodore Levin, Coppe van Urk
{mitcho, tedlevin, coppe}@alum.mit.edu
Austronesian Formal Linguistics Association 26
University of Western Ontario, May 2019

1 Introduction

Today we discuss the typology of Austronesian voice system languages and develop a new understanding of the attested range of variation.

(1) Sqiuliq Atayal:
   a. Cyx p-hapuy sehuy sa knobuy qu Yuraw.
      Aux AV:R-cook taro(acc) loc kitchen nom Yuraw
      ‘Yuraw cooks taro in the kitchen.’
      Actor Voice (AV)
   b. Puy-un na Yuraw qu sehuy.
      cook-pv gen Yuraw nom taro
      ‘Yuraw cooked taro.’
      Patient Voice (PV)
   c. Hpuy-an na Yuraw sehuy qu knobuy.
      cook-lv gen Yuraw taro(acc) nom kitchen
      ‘Yuraw cooks taro in the kitchen.’
      Locative Voice (LV)

(2) Agent extraction ⇒ AV:
   a. Ima wal m-aniq sehuy qas a?
      who past av-eat taro that
      ‘Who ate that taro?’
   b. *Ima wal niq-un qu sehuy qas a?
      what past av-extraction
      ‘What did Yuraw eat?’
   c. *Ima wal niq-un na Y?
      who past av-extraction
      ‘What did Yuraw do?’

(3) Patient extraction ⇒ PV:
   a. *Naru wal m-aniq qu Y?
      what past av-eat nom Y
      ‘Who ate that?’
   b. Naru wal niq-un na Y
      ‘What did Yuraw do?’

But voice system languages can also look pretty different...

(4) Balinese:
   a. Actor Voice (AV):
      Polisi ng-ejuk nyoman.
      Police av-arrest Nyoman
   b. Patient Voice (PV):
      Nyoman 0-ejuk polisi.
      Nyoman pv-arrest police
      ‘A policeman arrested N.’

This work is part of our ongoing work on the syntax of Austronesian-type voice systems within Austronesian and beyond. We thank Edith Aldridge, Mark Baker, TC Chen, Julie Legate, David Pesetsky, Masha Polinsky, Norvin Richards, and Lisa Travis for comments on this overarching project, and colleagues at LSA 2018. Errors are each other’s.

1 We use the terms “agent” and “actor” interchangeably. We apologize for any confusion.
2 In ELvU, we referred to the privileged argument as the “pivot”; in other work, it has also been called “focus,” “topic,” “trigger”. See Blust 2002, Ross and Teng 2005, Blust 2013 sec. 7.1 for discussion of these terms. Other additions to (7) are in [square brackets].

Previously at AFLA...³

1. The subject argument receives structural nominative case from a high functional head.
   • Nominative may be assigned to a nominal which already has a case value, overriding its realization (see Chen 2018 on Amis; Richards 2013; Levin 2017 for other examples).

2. The subject is uniquely positioned to be Agreed with by an $\overline{\Lambda}$-probe — e.g. is highest in the lower phase (Aldridge, 2004; Rackowski and Richards, 2005; Van Urk and Richards, 2015 and many others) — or always in a high $\overline{\Lambda}$-position $\bullet$ (Richards, 2000; Pearson, 2005).
   ▶ So the subject is associated with mixed A/$\overline{\Lambda}$ properties (Van Urk, 2015).
     • The functional head(s) associated with the subject’s A- and $\overline{\Lambda}$-properties may in fact be a single head, bundling features traditionally associated with C and T. We call this CT (Legate, 2011; Aldridge, 2017; Erlewine, 2018; see also Martinović, 2015).

3. Every nominal requires licensing (e.g. Case).

Note: “Licensing” is not morphological case value determination (Marantz, 1991).

(8) a. Actor Voice:

\[
\begin{array}{c}
\text{CTP} \\
\text{CT} \\
\text{vP} \\
\text{DP} \\
\text{agent subj} \\
\text{NOM} \\
\end{array}
\]

b. Non-Actor Voices:

\[
\begin{array}{c}
\text{CTP} \\
\text{CT} \\
\text{vP} \\
\text{DP} \\
\text{agent subj} \\
\text{NOM} \\
\end{array}
\]

The fact that non-subject agents are also at the phase edge predicts that — under the right circumstances — they may also be accessible for probing/attraction from above. See Erlewine and Levin 2018 and Erlewine and Lim 2018 for precisely such evidence.

³But on the relationship between voice morphology on the verb and the choice of subject argument, we currently feel less committed to thinking of this as “extraction marking” as we argued for in ELvU 2015, 2017.

3 Proposal

▶ How do non-subject arguments get licensed? The theory in §2 leaves this underspecified.

Three ways nominals can get licensed:

1 Structural Case — requires a particular structural configuration.
3 Prepositional/oblique case-insertion — adding a K/P head as a rescue strategy (Stowell, 1981; Halpert, 2012; Imanishi, 2014; Van Urk, 2015).

Austronesian voice system languages vary according to two parameters:

9 Accusative parameter:
   The language {does/does not} have structural accusative case. $\bullet$
   (Technically: $v$ {can/cannot} assign accusative case downward.)

10 Last-resort licensing parameter:
   If a DP lacks a source for structural licensing, it can be licensed...
   a. under linear adjacency with the verb, or $\bullet$
   b. by insertion of a case-marker (genitive). $\bullet$
4 A typology of voice systems

Our proposal in (9)–(10) predicts a four-way typology of voice system languages:

<table>
<thead>
<tr>
<th>Adjacency:</th>
<th>Toba Batak §4.1</th>
<th>Balinese §4.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-insertion:</td>
<td>Nanwang Puyuma §4.3</td>
<td>Katipul Puyuma §4.4</td>
</tr>
</tbody>
</table>

Here we concentrate on the behavior of core arguments: agents and themes.

4.1 No accusative + adjacency = Toba Batak (following Erlewine, 2018)

There is no source of structural licensing for non-subject arguments, so all non-subject core arguments must be licensed by adjacency with the verb.

(11) Non-subjects must be immediately post-verbal:
       \[av\text{-read yesterday book yesterday fn Pollak yesterday}\]
       ‘Pollak read a book yesterday.’
       \[rv\text{-read yesterday fn Pollak yesterday book yesterday}\]
       ‘Pollak read a book yesterday.’

See Erlewine 2018 for additional data which shows that post-verbal word order is free, with the one exception of the adjacency requirement on post-verbal non-subjects.

- This word order restriction applies symmetrically to non-subject agents and non-subject themes.
- The verb and post-verbal non-subject form a phonological or intonational unit for the purposes of main stress placement (Emmorey, 1984).
- Licensing by adjacency (Baker, 2014; Levin, 2015) requires head-head adjacency between V and the highest head of the nominal, imposing a directionality restriction. All nominals in Toba Batak are DPs and D heads are pre-nominal — e.g. si for personal names — so licensing by adjacency requires the DP to be immediately post-verbal; being immediately pre-verbal is insufficient.
- We predict that there are no ditransitives with three DPs in Toba Batak.
  - Why? Because licensing by adjacency with its directionality restriction can only license one non-subject DP.
  - This appears correct; all verbs elicited in Toba Batak have at most two DPs, with ditransitives taking an oblique (PP) goal; see Erlewine 2018: 678.

4.2 Accusative + adjacency = Balinese (following Levin, 2015)

Structural accusative can license non-subject themes, but not agents. Non-subject agents — but not non-subject themes — will require verb adjacency.

(12) Asymmetric adjacency restriction on non-subjects:
    a. Cicing ng-uber (ke jalan-e) siap-e.
       \[av\text{-chase into street-DEF chicken-DEF}\]
       ‘A dog chased the chicken (into the street).’ (Wechsler and Arka, 1998)
    b. Siap-e 0-uber (*ke jalan-e) cicing.
       \[av\text{-chase into street-DEF dog}\]
       ‘A dog chased the chicken (into the street).’ (Wechsler and Arka, 1998)

- Nominals licensed under adjacency in Balinese can be NPs or DPs. D heads are post-nominal (=e), so DPs with nominal domains cannot be licensed by adjacency:

(13) a. I Wayan 0-gugut cicing.
    \[av\text{-bite dog}\]
    ‘A dog bit Wayan.’
    b. *I Wayan 0-gugut cicing-e (ento).
    \[av\text{-bite dog-DEF (that)}\]
    ‘The dog bit Wayan.’

- In an NP, the highest head in the nominal functional sequence is N, so N must be adjacent to the verb. In this case, we yield a ban on pre-nominal adjectives:

(14) a. \{Liu\} cicing \{liu\} n-ugut Nyoman.
    \[many dogs many-DEF Nyoman\]
    ‘Many dogs bit Nyoman.’
    b. Nyoman 0-gugut \{liu\} cicing.
    \[av\text{-bite many-DEF dog}\]
    ‘Many dogs bit Nyoman.’

- But this is not a definiteness restriction: Pronouns and proper names can be licensed under adjacency, unlike with pseudo-noun incorporation in Niuean (Massam, 2001).

(15) a. Be-e 0-daar ida.
    \[av\text{-eat fish-DEF}\]
    ‘(S)he ate the fish.’
    b. Be-e 0-daar Nyoman.
    \[av\text{-eat fish-DEF Nyoman}\]
    ‘Nyoman ate the fish.’

We can derive this if pronouns and proper names occupy D0 (e.g. Postal, 1966; Longobardi, 1994; Elbourne, 2001) and lack an NP, satisfying head-head adjacency.
4.3 No accusative + case-insertion = Nanwang Puyuma

There is no source of structural licensing for non-subject arguments — just like in Toba Batak — so they will all be licensed with the same last-resort prepositional/oblique case.

(16) Non-subjects are all marked with the same case series:

<table>
<thead>
<tr>
<th>Case markers in Nanwang Puyuma (Teng, 2009: 827):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal name</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Subject</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Non-subject agent</td>
</tr>
<tr>
<td>Non-subject theme</td>
</tr>
</tbody>
</table>

• We call this case genitive as it is the same as that on possessors (18) (although possessor and non-subject bound pronominal series differ).

(17) Case markers in Nanwang Puyuma (Teng, 2009: 827):

4.4 Accusative + case-insertion = Katipul and Ulivelivek Puyuma

Structural accusative can license non-subject themes, but not agents — just like in Balinese. So non-subject agents alone will receive last-resort (genitive) case, distinct from accusative.

The more conservative Katipul and Ulivelivek dialects of Puyuma distinguish the case marking series for non-subject agents and non-subject themes:

(20) Case markers in Katipul Puyuma (Teng, 2009: 827):

<table>
<thead>
<tr>
<th>Case markers in Katipul Puyuma (Teng, 2009: 827):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal name</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Subject</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Non-subject agent</td>
</tr>
<tr>
<td>Non-subject theme</td>
</tr>
</tbody>
</table>

The case marker series for non-subject agents is again equivalent to that for possessors.

4.5 Summary

The four types of languages identified and highlighted here look on the surface to be quite different. For example...

• Puyuma has case markers; Toba Batak and Balinese do not.
• The Nanwang and Katipul varieties of Puyuma vary in the number of distinct case marking series.
• Balinese and Toba Batak both have restrictions on post-verbal word order, but they affect different arguments: Balinese requires non-subject agents alone to be immediately post-verbal, whereas Toba Batak requires all non-subject DPs to be immediately post-verbal.
• The two parameters for non-subject licensing in (9) and (10) help us productively understand and relate these superficially distinct classes of languages.

For example, the fact that Nanwang Puyuma and Toba Batak both treat all non-subjects in a symmetric fashion can be attributed to a lack of structural accusative — despite these two languages looking quite different at first glance.

4 However, a salient difference is that non-subject agents but not non-subject themes are obligatorily clitic-doubled on the verb, as seen in (16). See discussion in Erlewine and Levin 2018.
5 Teng glosses this as on.
6 This pattern could be derived through separate sources for the two case markers, but would then need to treat the surface forms as a case of accidental syncretism.
7 There is, however, a distinction reported for definite common noun possessors vs non-subject agents for Ulivelivek but not Katipul; see Teng 2009: 827.
5 Discussion and extensions

Our proposed parameters successfully describe the key behaviors of a range of different Austronesian voice system languages. By way of conclusion, we consider a couple extensions.

5.1 No accusative + case-insertion (Nanwang P.) + specificity DOM = Tagalog

Like in Nanwang Puyuma, non-subject agents and themes in Tagalog bear the same, genitive case marking in many cases:

(21) Non-subject agents and themes in genitive case: (Schachter, 1996)
   a. B<um>ili ang babae ng tela.
       av-bought nom woman gen cloth
       ‘The woman bought some cloth.’
   b. B<in>ili ang babae ng tela.
       rv-bought gen woman nom cloth
       ‘A/The woman bought the cloth.’

   ▶ But notice that there is an interaction with the theme’s specificity, with the genitive non-subject theme in (21-a) necessarily being non-specific. Instead, specific non-subject themes are oblique, as in (22). 8
   (Schachter and Otanes, 1972; McFarland, 1978; Sabbagh, 2016; a.o.)

(22) Sino ang b<um>aril sa ibon?
    who nom av-shot obl bird
    ‘Who shot the bird?’

   ▶ Tagalog has no accusative for non-specific non-subject themes, but can assign (obl) case to specific non-subject themes (see e.g. Sabbagh, 2016).
   – With non-specific themes, Tagalog resembles Nanwang Puyuma; with specific themes, Tagalog resembles Katipul and Ulivelivek Puyuma!

[The unavailability of the specific theme interpretation in (22-a) could be described as due to the availability of the competing form in (21-b). When the PV form is blocked, for example by the extraction restriction, a specific theme interpretation becomes available:]

(i) Sino ang b<um>aril ng ibon?
    who nom av-shot gen bird
    ‘Who shot a/the bird?’ (McFarland, 1978: 149)

5.2 Accusative + licensing by adjacency (Balinese) + -ny = Malagasy

(23) Non-subject agents must be adjacent to the verb: (Pearson, 2005)
   a. Nohanin’ ny gidro haingana ny voankazo omaly.
      pv.rv.eat det lemur quickly det fruit yesterday
      ‘The lemur ate the fruit quickly yesterday.’
   b. *Nohanin(a) haingana ny gidro ny voankazo omaly.
   c. *Nohanin(a) omaly ny gidro haingana ny voankazo.

   The post-verbal non-subject agent forms a tight phonological unit with the verb, which for pronouns and names results in being written as a single word, with word-internal phonological processes applying:

      pv.kill.1s [kill-pv-LNK 1sg] with-det knife det chicken
      ‘I am killing the chickens with the knife.’
      pv.kill-Ramatoa [kill-pv-LNK Ramatoa] with-det knife det chicken
      ‘Ramatoa is killing the chickens with the knife.’

   ▶ But notice that a linker (-ny) appears in these contexts as well. What is -ny doing?
   – One idea: Malagasy D heads expect to be valued for a morphological case feature specification. But in the absence of such a specification, the head is realized as =ny.
   – It is not unusual for the absence of feature specifications — or the most unmarked feature specification — to result in overt morphology: Consider English third singular -s.

5.3 [Your language here]

   ▶ Where does your language fit in? Thank you!
References


