Word order and disambiguation in Pangasinan

Joey Lim
joey.lim@nus.edu.sg

Michael Yoshitaka Erlewine
mitcho@nus.edu.sg

SICOGG 22, August 2020

NUS
National University of Singapore

Introduction

- There is a well-known typological trade-off between **word order flexibility** and **case and/or agreement** (Sinnemäki 2008, Fedzechkina et al. 2017, a.o.).
    - A simple, functionalist explanation: the mapping between arguments and verbs should be relatively unambiguous.
    - This "choice" is often conceived of as a *language-level* parameter.

Introduction

Today, we investigate the interaction of word order, case, and agreement in **Pangasinan** (Austronesian; Philippines).

- Descriptively, the functional pressure of disambiguation is *active in an individual grammar*:
  - **Post-verbal word order is free except when two arguments are *formally indistinguishable*, i.e. by formal features on the arguments.**
- We can explain these facts through a particular **feature-driven approach to scrambling.**

Data

- Data from elicitation with three native speakers of Pangasinan currently residing in Singapore, with some variation.
- The crucial judgments here represent the internally consistent behavior of a speaker from Lingayen, the capital of Pangasinan.

§2 Voice and case in Pangasinan

Voice system in Pangasinan

Pangasinan is predicate-initial and exhibits a **"voice system"**:

- In each clause, one argument is selected as the **"pivot"** and bears nominative case.
- The choice of pivot is reflected by morphology on the verb.
- Non-pivot arguments bear other cases, e.g. genitive or oblique.

Voice system in Pangasinan

(1) **Actor Voice (AV):**

Nan-sulat may laki la liham.
PFV.AV-write NOM man GEN letter
'The man wrote a letter.'

(2) **Patient Voice (PV):**

In-sulat may liham la laki.
PFV.PV-write NOM letter GEN man
'The man wrote a letter.'

There are also other, peripheral Voices (Locative, Benefactive, etc.), but we concentrate on Actor Voice vs Patient Voice here.

Case markers in Pangasinan

Genitive: *la*

Nominative:

- *si* on proper names;
- *may* and *su* on common nouns
  (*may* is singular; *(i)ra-may* is plural)

May and *su* generally appear to be in free variation:

(3) Nan-sulat **may/su** laki la liham.
PFV.AV-write NOM man GEN letter
'The man wrote a letter.'

However, their behavior will come apart later.

Double nominative constructions

- Pangasinan allows for Non-Actor Voice clauses where **both the pivot and non-pivot agent receive nominative case:**

(4) In-sulat =to **may** laki **may** liham.
PFV.PV-write =3SG.GEN NOM man NOM letter
'The man wrote a letter.'

This pattern is not attested, to our knowledge, in sister languages such as Tagalog and Bikol.

Properties of the double nominative

1. Double nominatives can only appear in Non-Actor Voices:

(5) a. * Nan-sulat (=to) **may** laki **may** liham.
PFV.AV-write =3SG.GEN NOM man NOM letter
'The man wrote a letter.' AV

b. In-sulat =to **may** laki **may** liham.
PFV.PV-write =3SG.GEN NOM man NOM letter
'The man wrote a letter.' PV

Other voices such as Benefactive Voice pattern with Patient Voice.

Properties of the double nominative

2. The agent must be clitic-doubled by a genitive pronoun:

(6) a. Lu-luto-en ***(=to)** may lakin ugaw ira-may sira.
IMPF-cook-PV =3SG.GEN NOM male child PL-NOM fish
'The boy is cooking the fishes.'

b. Lu-luto-en ***(=da)** ra-may lakin ugaw may sira.
IMPF-cook-PV =3PL.GEN PL-NOM male child NOM fish
'The boys are cooking the fish.'

These pronouns are second-position clitics.

Properties of the double nominative

3. The pivot can be marked *su* or *may*; the non-pivot agent must be marked *may*. There are no double-*su* clauses:

(7) * S<in>aliw =to **su** lakin ugaw **su** aso.
PV.PFV-buy =3SG.GEN NOM male child NOM dog
'The boy bought the dog.'

In *su–may* combinations, it is clear which argument is which, regardless of word order:

(8) S<in>aliw =to **su** lakin ugaw **may** aso.
PV.PFV-buy =3SG.GEN NOM male child NOM dog
a. * 'The boy bought the dog.'
b. 'The dog bought the boy.'

§3 Word order and disambiguation

Word order and disambiguation

- Postverbal word order in Pangasinan is indeed free in conventional, non-double-nominative clauses, just as has been described for sister languages such as Tagalog (Kroeger 1991/1993, Richards 1993).
- However, in double nominative clauses, **a word order restriction emerges *specifically when two arguments are formally indistinguishable.***

Word order freedom

In non-double-nominative clauses, postverbal word order is free.

Arguments can be scrambled without affecting the interpretation:

(9) Man-lu-luto { **may** laki **la** sira / **la** sira **may** laki }.
AV-IMPF-cook NOM male GEN fish GEN fish NOM male
'The boy is cooking the fish.'

(10) Lu-luto-en { **la** laki **may** sira / **may** sira **la** laki }.
IMPF-cook-PV GEN male NOM fish NOM fish GEN male
'The boy is cooking the fish.'

Word order in the double nominative

In the double nominative, the two arguments are more easily confused: they can both appear with the same marking, *may–may*.

- **When two arguments are confusable (e.g. *may–may*), their word order is restricted to "agent < pivot" order.**
  - In practice, there are many ways in which two arguments can be formally distinguished, and then the word order is again free.

(This is the internally consistent behavior of one speaker's grammar. We comment on other speakers' judgements at the end.)

Confusability ⇒ word order rigidity

Consider the double nominative with "*may* girl *may* boy" order:

(11) Pinu-niti =to **may** bien ugaw **may** lakin ugaw.
PV.PFV-hit =3SG.GEN NOM female child NOM male child

a. ag < pivot (th): 'The girl hit the boy.'

b. * pivot (th) < ag: 'The boy hit the girl.'

Confusability ⇒ word order rigidity

This is so even if the arguments differ in animacy:

(12) Lu-luto-en =to may sira may lakin ugaw.
IMPF-cook-PV =3SG.GEN NOM fish NOM male child

a. # ag < pivot (th): 'The fish is cooking the boy.'

b. * pivot (th) < ag: 'The boy is cooking the fish.'

Confusability ⇒ word order rigidity

…or if the arguments should be distinguishable by world knowledge:

(13) A-nengneng =to may bulag ya laki may bie.
POT.PV-see =3SG.GEN NOM blind ATTR male NOM female

a. # ag < pivot (th): ‘The blind man saw the woman.’

b. * pivot (th) < ag: ‘The woman saw the blind man.’

Distinguishability ⇒ word order flexibility

However, there are also many double nominative clauses with free word order between the pivot and agent. This includes cases where:

- the pivot and agent take different nominatives (*su* vs *may*)
  (But this isn't just about surface form: proper names with *si* pattern with *may*, so *si*–*may* clauses are restricted to "agent < pivot" order.)
- the pivot and agent differ in number, leading to disambiguation by φ-agreement;
- one argument is a reflexive (shown later).

Distinguishability ⇒ word order flexibility

Recall that pivots can be marked *su* or *may*, whereas non-pivot agents are marked *may*. When the pivot is marked with *su*, their relative word order is free:

(14) a. Pinu-niti =to **su** bien ugaw **may** lakin ugaw.
PV.PFV-hit =3SG.GEN NOM female child NOM male child
pivot (th) **ag**
'The boy hit the girl.' / *'The girl hit the boy.'

b. Pinu-niti =to **may** lakin ugaw **su** bien ugaw.
PV.PFV-hit =3SG.GEN NOM male child NOM female child
ag **pivot (th)**
'The boy hit the girl.' / *'The girl hit the boy.'

Distinguishability ⇒ word order flexibility

Word order is also free when φ features of the two arguments differ:

(15) 3sg agent / 3pl theme:

a. Lu-luto-en =**to** **may** lakin ugaw **ira-may** sira.
IMPF-cook-PV =3SG.GEN NOM male child PL-NOM fish
ag **pivot (th)**
'The boy is cooking the fishes.'

b. Lu-luto-en =**to** **ra-may** sira **may** lakin ugaw.
IMPF-cook-PV =3SG.GEN PL-NOM fish NOM male child
pivot (th) **ag**
'The boy is cooking the fishes.'

Recall that the clitic pronoun (here: third-singular) always cross-references the agent, disambiguating.

Distinguishability ⇒ word order flexibility

The examples in (16) differ from (15) only in the clitic pronoun, now third-plural *=da*:

(16) 3pl agent / 3sg theme:

a. Lu-luto-en **=da** ra-may lakin ugaw may sira.
IMPF-cook-PV =3PL.GEN PL-NOM male child NOM fish
ag **pivot (th)**
'The boys are cooking the fish.'

b. Lu-luto-en **=da** may sira ira-may lakin ugaw.
IMPF-cook-PV =3PL.GEN NOM fish PL-NOM male child
pivot (th) **ag**
'The boys are cooking the fish.'

Confusability again ⇒ word order rigidity again

With two third-plural arguments, word order is again restricted:

(17) 3pl agent / 3pl theme:

a. Lu-luto-en =da ra-may laki ira-may sira.
IMPF-cook-PV =3PL.GEN PL-NOM male PL-NOM fish
ag **pivot (th)**

i. 'The boys are cooking the fishes.'

ii. * 'The fishes are cooking the boys.'

b. # Lu-luto-en =da ra-may sira ira-may laki.
IMPF-cook-PV =3PL.GEN PL-NOM fish PL-NOM male
ag **pivot (th)**

i. # 'The fishes are cooking the boys.'

ii. * 'The boys are cooking the fishes.'

Summary

- Thus, word order is fixed precisely when the two arguments cannot be *formally distinguished*:
    - by case marking, φ-features, or reflexive status (below)
- When they can be distinguished by different case markers, different φ features or reflexive status, their word order is free and arguments can be scrambled.

§4 Proposal

Proposal

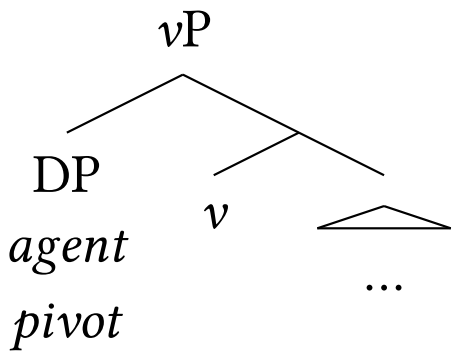
Our analysis, in three parts:

1. A basic theory for Philippine voice systems
   (ELvU = Erlewine, Levin, and Van Urk 2017, 2020, in prep)
2. Extension to double nominatives
3. A story for scrambling

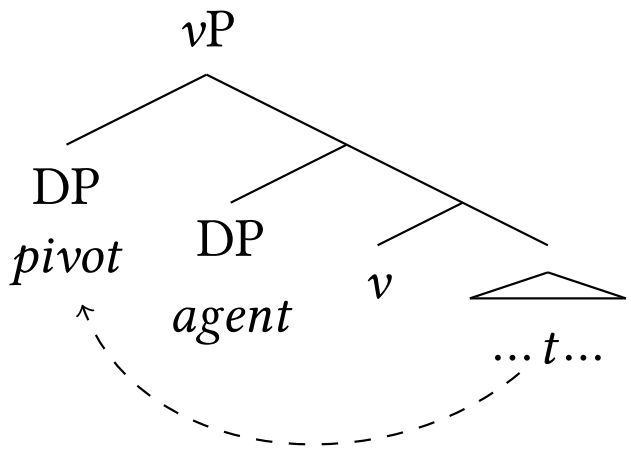
A basic theory for voice systems

- Following Aldridge 2004, Rackowski and Richards 2005 a.o., the pivot is the highest argument of the lower (*v*P) phase:

(18) a. Actor Voice:



b. Non-Actor Voices:

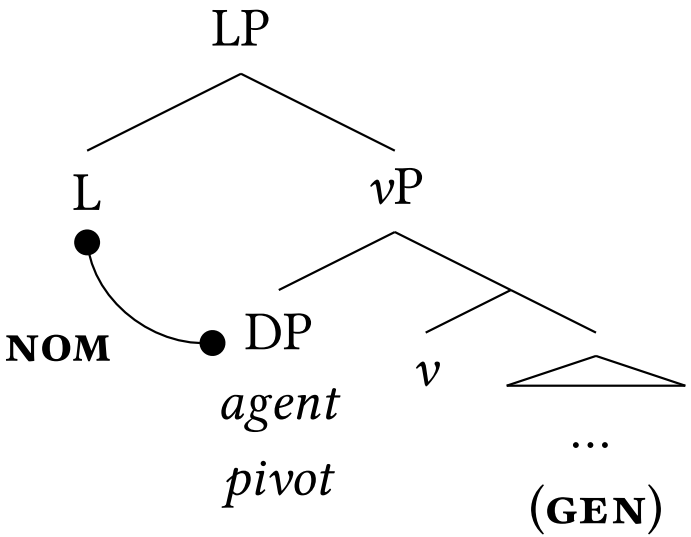


This pivot movement in NAV is a covert movement.

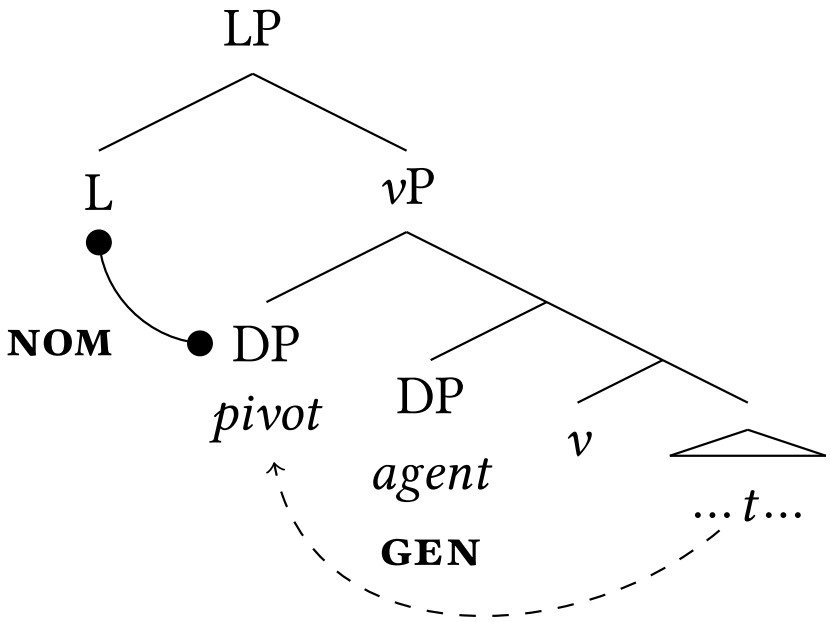
A basic theory for voice systems

- The pivot receives **structural nominative** (ELvU 2015, 2017) from a Licensing head L above *v*P (Halpert 2016);
- Nominals in *v*P without oblique or structural case are realized with **default genitive** (Erlewine and Lim 2019, ELvU 2020).

(19) a. Actor Voice:



b. Non-Actor Voices:



A basic theory for voice systems

- The verbal complex head-moves up, to be leftmost.
- Without *scrambling*, we yield "agent < …" order in all cases:
  - AV: "V NOM=agent/pivot … (GEN=…) …"
  - NAV: "V GEN=agent … NOM=pivot …"

  See Guilfoyle, Hung, and Travis 1992, Sells 2000 a.o. for evidence that "V agent … pivot …" is at least a preferred and possibly base order across various Philippine languages.

Extension to double nominatives

For double nominatives in Pangasinan (for all our speakers) — unattested in Tagalog and Bikol — we propose **two probes on L:**

(20)
i. obligatory; targets the closest DP and assigns structural nominative (*may* or *su*).
ii. optional, probes second; targets the *next* closest DP and
   (a) copies its φ-features, to be realized as a genitive clitic pronoun, and
   (b) assigns it restricted nominative case (*may*).

Extension to double nominatives

- Probe (i) necessarily targets the pivot, the highest DP in *v*P, explaining the pivot's appearance with *su* or *may*. Probe (ii) optionally probes after (i), targeting the agent.
- The licensing probes on L are subject to Phase Impenetrability. In AV clauses, there is only one DP at the *v*P phase edge, which receives nominative (i). There is no effect of optional probing by (ii), explaining the lack of double nominative AV clauses.

Scrambling

- Suppose that **scrambling is feature-driven** (see e.g. Grewendorf and Sabel 1999) — in particular, by **optionally building probes for arbitrary feature bundles** — and subject to locality (Attract Closest/Minimal Link Condition).

(21) [PROBE:**Y**] … [$_{v\text{P}}$ … α[X] … β[**Y**] … ⇒ "β … α … ___ …"

(Phases must not be relevant for this scrambling.)

Scrambling

Q: What if α and β are *featurally identical*?

A: If the default is "$\alpha < \beta$," **we predict "$\boldsymbol{\beta < \alpha}$" order to be underivable by scrambling!**

Scrambling

Q: What features can these scrambling probes target?

A: Based on patterns above, they're *morphosyntactic* features:

- CASE features (Deal 2017), assuming $\text{NOM}_{su} \neq \text{NOM}_{may}$;
- sub-features of φ, e.g. [SG], [PL];
- a feature on reflexives [REFL] (see below).

Scrambling

- Crucially, for the judgments presented above: **scrambling cannot involve probing for optional, Ā-features!**
  - If probing for an optional Ā-features is possible, we predict any phrase to be scramble-able: just add feature [SCR] to some phrase and build [PROBE:SCR] above (see e.g. Sauerland 1999, Müller 1998, 2002).

A note on dialectal/ideolectal variation

We note however that the rigid word order judgments of (11–13, 17), which our theory here accounts for, are not shared by two other speakers we have worked with:

- These two other speakers *do* allow the use of world knowledge to disambiguate arguments. In cases where world knowledge does not disambiguate, sentences may be ambiguous, although there is still an "agent < pivot" preference.
- We suggest that these other speakers *do* make use of an optional [SCR] feature for scrambling, allowing for the scrambling of likes.

Scrambling: A puzzle

Scrambling is Ā-movement, as has been claimed for Tagalog (Richards 1993):

(22) a. Agents bind themes:

Aka-nengneng may laki$_i$ ed *sarili=to*$_i$.
PFV.AV-see NOM male OBL self=3SG.GEN
'The boy saw himself.'

b. Unaffected by scrambling:

Aka-nengneng ed *sarili=to*$_i$ may laki$_i$ ___.
PFV.AV-see OBL self=3SG.GEN NOM male
'The boy saw himself.'

Scrambling: A puzzle

- But scrambling (for our first speaker) clearly cannot be due to probing for an optional Ā-feature!
- This challenges Van Urk's (2015) featural view of the A/Ā-distinction: targeting obligatory features yields A-movement, targeting optional features yields Ā-movement.

Scrambling: A second puzzle

Interestingly, with *may–may* double nominatives,

(23) a. "*may* boy *may himself*":

A-nengneng =to may laki$_i$ may *sarili=to*$_i$.
POT.PV-see =3SG.GEN NOM male NOM self=3SG.GEN
'The boy saw himself.'

b. Unaffected by scrambling!!

A-nengneng =to may *sarili=to*$_i$ may laki$_i$ ___.
POT.PV-see =3SG.GEN OBL self=3SG.GEN NOM male
'The boy saw himself.'

Assuming that this is the agent binding the theme again, and a theme cannot bind an agent, (23b) suggests that reflexives ([REFL]) can also be probed for in scrambling.

§5 Conclusion

Conclusion

- Languages appear to choose between case/agreement and rigid word order, to ensure unambiguous argument mappings.
- Today, we showed that **such a trade-off is visible *within a single grammar*:**
  - The Pangasinan double nominative results in clauses with multiple *formally indistinguishable* arguments (e.g. *may–may*).
  - In such cases, for one speaker, **scrambling is restricted.**

Conclusion

- We propose an analysis based on **scrambling as a feature-driven movement.**
  - For one speaker, scrambling can only probe for *morphosyntactic features* such as case/φ features, not an optional Ā-feature, explaining the word order restrictions.
  - This leaves us with a puzzle for the Ā-properties of scrambling.

Future directions

In further work, we also intend to investigate…

- interaction with preverbal fronting;
- other types of nominals;
- other A/Ā-diagnostics.

We welcome comments and suggestions for this ongoing project!

Thank you!

Thank you! Questions?

We thank our speakers and NUS syntax/semantics lab members, especially Kenyon Branan, for discussion.

References I

Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.

Deal, Amy Rose. 2017. Syntactic ergativity as case discrimination. In *Proceedings of WCCFL 34*, ed. Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel, and Edward J. Rubin, 141–150.

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe Van Urk. 2015. What makes a voice system? On the relationship between voice marking and case. In *AFLA 21: The Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association*, ed. Amber Camp, Yuko Otsuka, Claire Stabile, and Nozomi Tanaka, 51–68. Asia-Pacific Linguistics.

References II

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe Van Urk. 2017. Ergativity and Austronesian-type voice systems. In *Oxford Handbook of Ergativity*, ed. Jessica Coon, Diane Massam, and Lisa deMena Travis, 373–396. Oxford University Press.

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe Van Urk. 2020. The typology of nominal licensing in Austronesian voice system languages. In *Proceedings of AFLA 26*, 71–87. URL `https://ling.auf.net/lingbuzz/004879/current.pdf`.

Erlewine, Michael Yoshitaka, and Cheryl Lim. 2019. Bikol clefts and topics and the Austronesian extraction restriction. URL `https://ling.auf.net/lingbuzz/004181/current.pdf`, manuscript, National University of Singapore.

References III

Fedzechkina, Maryia, Elissa L. Newport, and T. Florian Jaeger. 2017. Balancing effort and information transmission during language acquisition: Evidence from word order and case marking. *Cognitive Science* 41:416–446.

Grewendorf, Günther, and Joachim Sabel. 1999. Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language & Linguistic Theory* 17:1–65.

Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory* 10:375–414.

Halpert, Claire. 2016. *Argument licensing and agreement*. Oxford University Press.

Kroeger, Paul R. 1991/1993. Phrase structure and grammatical relations in Tagalog. Doctoral Dissertation, Stanford University.

References IV

Müller, Gereon. 1998. *Incomplete category fronting*.

Müller, Gereon. 2002. Two types of remnant movement. In *Dimensions of movement: From features to remnants*, ed. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, and Hans-Martin Gärtner, 209–241. John Benjamins.

Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565–599.

Richards, Norvin. 1993. Tagalog and the typology of scrambling. Honors thesis, Cornell University.

Sauerland, Uli. 1999. Erasability and interpretation. *Syntax* 2:161–188.

Sells, Peter. 2000. Raising and the order of clausal constituents in the Philippine languages. In *Formal issues in Austronesian linguistics*, ed. Ileana Paul, Vivianne Phillips, and Lisa Travis, 117–143. Springer.

References V

Sinnemäki, Kaius. 2008. Complexity trade-offs in core argument marking. In *Language complexity*, ed. Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson, 67–88. John Benjamins.

Van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A Dinka Bor case study. Doctoral Dissertation, Massachusetts Institute of Technology.