Object Voice in Budai Rukai*

Chun-Hung Shih | University of Pennsylvania

AFLA 32 | June 3, 2025

1 Introduction

- (Western) Austronesian languages can be broadly categorized into Philippine-type languages and Indonesian-type languages (Wolff 1996, Himmelmann 2001, V. Chen and McDonnell 2019).
- Philippine-type languages have four 'voices,' which are the actor voice (AV), patient voice (PV), locative voice (LV), and circumstantial voice (CV), as seen in Paiwan.[1] Most Formosan languages, Austronesian languages spoken in Taiwan, follow this type.

(1) a. Q<**m**>alup **a** **caucau** tu vavuy i gadu tua vuluq.
<**AV**>hunt **NOM man** OBL pig LOC mountain OBL spear
'The man hunts wild pigs in the mountains with a spear.' (Ferrell 1979:202)

b. Qalup**-en** nua caucau **a** **vavuy** i gadu tua vuluq.
hunt**-PV** GEN man **NOM pig** LOC mountain OBL spear
'The man hunts wild pigs in the mountains with a spear.' (ibid)

c. Qalup**-an** nua caucau tua vavuy **a** **gadu** tua vuluq.
hunt**-LV** GEN man OBL pig **NOM mountain** OBL spear
'The man hunts wild pigs in the mountains with a spear.' (ibid)

d. **Si-**qalup nua caucau tua vavuy i gadu **a** **vuluq**.
CV-hunt GEN man OBL pig LOC mountain **NOM spear**
'The man hunts wild pigs in the mountains with a spear.' (ibid)

- Indonesian-type languages are characterized with three voices, which are the active voice (AV), object voice (OV), and passive voice, as seen in Indonesian. OV is typically bare.

(2) a. **Kami** tidak akan **mem-baca** buku ini.
1PL not will **AV-read** book this
'We will not read this book.' (Cole et al. 2008:1512)

b. **Buku ini** tidak akan kami **baca**.
book this not will 1PL **read**
'This book will not be read by us.' (ibid)

c. **Buku ini** tidak akan **di-baca** oleh Siti.
book this not will **PASS-read** by Siti
'This book will not be read by Siti.' (ibid)

*I'd like to express my gratitude to Julie Anne Legate, Matt Hewett, and Ka Fai Yip for their advice; to Sulryape ka taiAluladen and Legeane ka taiKasepelane for sharing their linguistic knowledge with me; to the audiences in Penn Synax Lab and Penn Meaning Lab where some parts of this project have been presented. The data presented here were mainly worked out with Sulryape and some of them were confirmed by Legeane. All errors, if any, are mine.

[1] It has to be noted that the terms 'voice', 'nominative', and 'subject' are adopted from descriptive linguistics in Austronesian literature and not used in a theoretical way. For the expository reason, I will adopt these terms for descriptions in my data as well.

- Budai Rukai is known for its active-passive (3a-b) distinction among Formosan languages (C.-F. Chen 1999, 2008, Zeitoun 2000, 2023, Sung 2011, Shih 2025). Note that the active voice (AV) is marked with an overt realis marker, without a dedicated AV morpheme.

(3) a. Wa-pana=aku ku babuy.
REAL-hunt=1SG.NOM OBL boar
'I hunted a boar.'

b. **Ki**-a-pana nakwane ka babuy.
PASS-REAL-hunt 1SG.OBL NOM boar
'That boar was hunted by me.'

- Morphologically, it also shares the object voice (OV; bare verb) (4), similar to Indonesian-type voice system;[2] it further exhibits the patient voice (PV) (5a-b), locative voice (LV) (5c), and circumstantial voice (CV) (5d), similar to the Philippine-type voice system.

(4) a. Pana=li **ka babuy**.
hunt=1SG.GEN **NOM boar**
'That boar is one that I hunted.'

(5) a. **Ni**-pana-ane nakwane **ka babuy**.
PV.REAL-hunt-NMLZ 1SG.OBL **NOM boar**
'That boar is one that was HUNTED by me.'

b. **A**-pana-ane=li **ka babuy**.
PV.IRR-hunt-NMLZ=1SG.GEN **NOM boar**
'That boar is one that I will hunt.'

c. **Ta**-pana-ane=li ki babuy **ka lregelrege**.
LV-hunt-NMLZ=1SG.GEN OBL boar **NOM mountain**
'That mountain is where I hunted the boar.'

d. **Sa**-pana-ane=li ki babuy **ka kwange**.
CV-hunt-NMLZ=1SG.GEN OBL boar **NOM pistol**
'That pistol is what I will use to hunt the boar.'

- The OV construction has been described and documented mostly in questions and relativization (Guo 1979, C.-F. Chen 1999, Zeitoun 2000, Sung 2011), but less attention is given in simple declaratives and its structure.

- **Goal**: to investigate the object voice in Budai Rukai, in terms of the clausal structure and the grammatical relations of agents and themes (semantics and information structure set aside)

- **Claims**: In Budai Rukai:

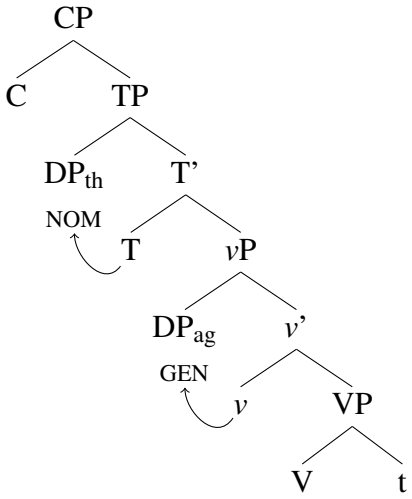
  - OV clauses show both verbal and nominal properties, and OV predicates are nominalized above MoodP.
  - OV agents are arguments and receive structural genitive Case from the nominalizer though being syntactically inert.
  - OV themes have a dual status in both Spec,PredP and Comp,VP within nominalization.

[2] C.-F. Chen 2008 uses the term OV for another construction in Budai Rukai where the verb is marked by *ta-...-ane*, and he proposes a view of nominalization to it with a main focus on semantics. Note that the OV discussed here (with bare verbs) is different from the OV in C.-F. Chen 2008, but I share his view on nominalization and explore syntax of nominalization. The reason that I use OV for the bare verb construction is that morphologically the bare OV seems to pattern with bare OV in (some) Indonesian-type languages.

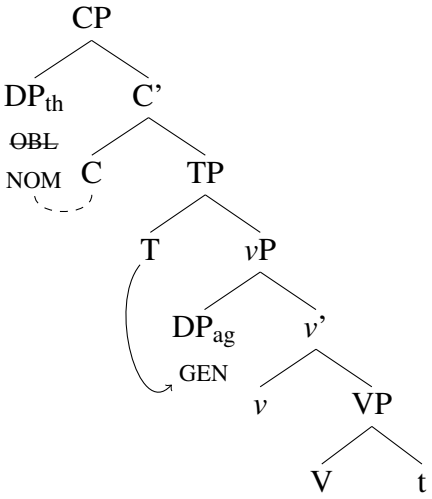
2 Theoretical analyses and main proposal

- Theoretically, three main analyses have been proposed for PV and/or OV.
  - **ergative approach**: nominative arguments are grammatical subjects and genitive agents are ergative, with PV and OV being voice morphemes
    (Aldridge 2004, 2008, 2012, H. Chang 2015, Nie 2020 for PV in Philippine-type languages; Arka 2003, 2008, Cole and Hermon 2005, Cole et al. 2008, Aldridge 2008, Legate 2014 for OV in Indonesian-type languages)
  - **agreement approach**: 'nominative' arguments are A'-elements of an overwritten topic case and 'genitive' agents have a structural Case, with PV and OV being agreement morphemes (Richards 2000, Pearson 2001, 2005, Rackowski 2002, Rackowski and Richards 2005, M. Chang 2004, V. Chen 2017, Erlewine et al. 2017 for PV in Philippine-type languages; Patrianto 2023 for OV in (some) Indonesian-type languages)

(6) ergative approach (linear set aside)

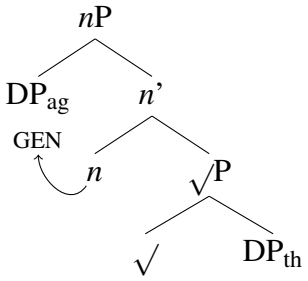


(7) agreement approach (linear set aside)

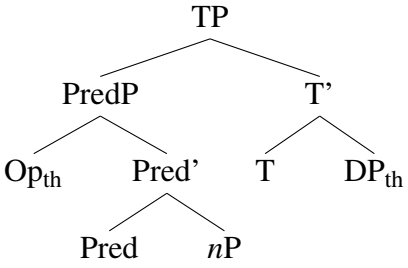


  - **nominalization approach**: nominative arguments are grammatical subjects and genitive agents are genitive in nominalization, with PV being nominalizers and predicates being nominal (Kaufman 2009, 2017)

(8)



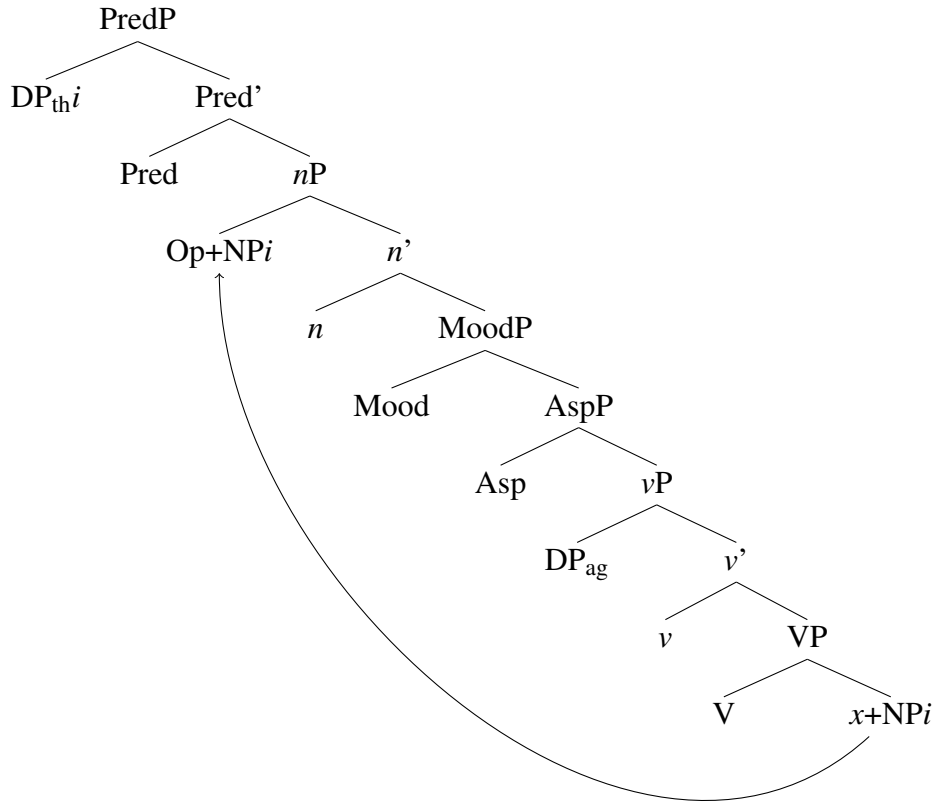
(9)



- **Proposal**: OV in Budai Rukai is nominalized, thus different from the ergative/agreement approaches (Rackowski 2002, Aldridge 2004, Legate 2014, V. Chen 2017, Patrianto 2023, a.o.). The nominalization is above MoodP, also different from root nominalization in the nominalization approach (Kaufman 2009, 2017).
  - **Matrix clause**: The nominalized structure serves as the predicate, as in the complement of $Pred^0$, whereas the theme argument serves as the subject, as base-generated in Spec,PredP.
  - **Within nominalization**: A full verbal projection is present with AspP and MoodP. The agent is introduced in Spec,*v*P, while the theme argument is in Comp,VP and moves up to Spec,*n*P to be identified with the theme in Spec,PredP. The theme is internally complex in

having an operator/variable with an NP-complement which is PF-deleted but LF-interpreted for reconstruction effects (see Salzmann 2019).

(10)



3 Predicates

- **Argument**: OV predicates show both verbal and nominal properties, suggestive of some verbal projections and nominalization. The nominalization is below TP and above MoodP.

3.1 Verbal properties

- **Adverbs**: Adverbs can show up to modify AV verbs (see Shih 2022) and also OV predicates.

(11) a. Wa-langay **kibubuli** ka didiwsa ka ama-Ø.
REAL-buy **stealthily** OBL car NOM father-1SG.GEN
'My father bought the car stealthily.'

b. Ma-dalame **turamuru** ki Lavurase ka Muni.
STAT.REAL-like **very** OBL Lavurase NOM Muni
'Muni likes Lavurase very much.'

(12) a. Langay **kibubuli** ki ama-Ø ka didiwsa.
buy **stealthily** GEN father-1SG.GEN NOM car
'That car is one that my father bought stealthily.'

b. Ka-dalame **turamuru** ki Muni ka Lavurase.
STAT-like **very** GEN Muni NOM Lavurase
'Lavurase is one that Muni likes very much.'

- **Progressive aspect**: AV verbs can have the progressive aspect, while nominal predicates need a verbalizer to form the progressive. OV predicates can have the progressive aspect without an additional morpheme, just like AV verbs.

(13) a. Wa-lapu-lapu=aku ka kaange.
REAL-RED-keep=1SG.NOM OBL fish
'I am keeping a fish.'

b. *(A-)si<nsi>nsi=aku.
*(A-)<RED>teacher=1SG.NOM
'I am being a teacher.'

(14) a. Lapu-lapu=li ka kaange.
RED-keep=1SG.GEN NOM fish
'That fish is one that I am keeping.'

3.2 Nominal properties

- **Negation**: In negation, the morpheme *ka-* is required in nominal predication but not in AV verbs. The morpheme *ka-* is required with OV predicates in negation, just like nominal predicates.

(15) a. Kai=su *(ka) sinsi.
NEG=2SG.NOM *(KA) teacher
'You are not a teacher.'

b. Kai=naku (*ka) wa-lrumay ki Lavurase.
NEG=1SG.NOM (*KA) REAL-hit OBL Lavurase
'I didn't hit Lavurase.'

(16) a. Kai *(ka) kirikiri ki ina-Ø ka damay.
NEG *(KA) stir.fry GEN mother-1SG.GEN NOM dish
'That dish is not one that my mother stir-fried.'

- **Plural morphology**: The plural morpheme is expressed by the prefix *la-* in Budai Rukai, which can show up on both nominal arguments and nominal predicates. The plural morpheme can also show up on OV predicates, which suggests that OV predicates are nominalized.[3]

(17) a. Wa-kela ki cekele kay **la**-vavalake.
REAL-come OBL tribe this **PL**-child
'These children came back to the tribe.'

b. **La**-sinsi=nay.
PL-teacher=1PL.EXCL.NOM
'We are teachers.'

(18) a. Ka la-karadrare, na-**la**-darepe=li.
NOM PL-bamboo.basket PAST-**PL**-weave=1SG.GEN
'Those bamboo baskets were ones that I wove.'

3.3 Height of nominalization: tense/mood and plural morphology

- **Morpheme order in PV**: With plural morphology, the plural morpheme follows the tense morphemes but precedes the (ir)realis morphemes. This indicates that plural morphology is lower than tense morphology but higher than (ir)realis morphology (Mirror Principle, Baker 1985). The nominalization is thus lower than TP and higher than MoodP.

[3]Note that the plural morphology on OV predicates is not necessary and often omitted in natural elicitation.

(19) a. Ka la-balruku, na-**la**-ni-darep-ane nakwane.
NOM PL-bamboo.tray PAST-**PL**-PV.REAL-weave-NMLZ 1SG.OBL
'Those bamboo trays were ones that I wove.'

b. Kavay la-kadangeane, lri-ka-**la**-a-langadh-ane=li.
this PL-land FUT-KA-**PL**-PV.IRR-buy-NMLZ=1SG.GEN
'These lands will be ones that I will buy.'

- **Morpheme order in OV**: The plural morpheme also follows the tense morpheme.[4] Despite no phonetic content, it can be assumed in parallel that a covert realis morpheme appears after the plural morpheme in OV and lower than nominalization.

- As lower, the covert realis morpheme instead of the overt one in AV (*wa-*) can thus be selected by the OV nominalizer.

(20) a. Ka la-karadrare, na-**la**-Ø-darepe=li.
NOM PV-bamboo.basket PAST-**PL**-OV.REAL-weave=1SG.GEN
'Those bamboo baskets were ones that I wove.'

b. Ka la-tawbubane, na-**la**-Ø-pia=li.
NOM PL-sheath PAST-**PL**-OV.REAL-do=1SG.GEN
'Those sheaths were ones that I made.'

3.4 Summary

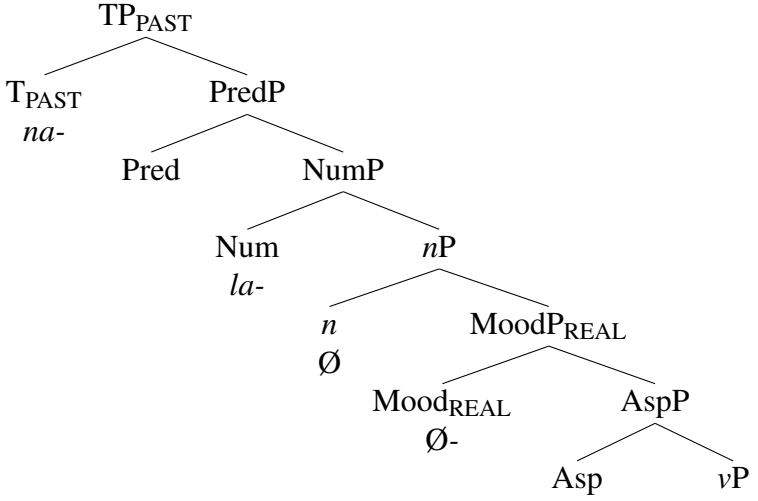
- OV predicates show nominalization where the structure is nominalized (1) above low adverbs, progressive aspect, and presumably realis mood, and (2) below tense and negation. This suggests that the nominalization is right above MoodP rather than at the root (Kaufman 2009, 2017).

- The height of nominalization follows the functional hierarchy proposed by Cinque (1999).

(21) $\text{Mood}_{\text{speech act}} > \text{Mood}_{\text{evaluative}} > \text{Mood}_{\text{evidential}} > \text{Mod}_{\text{epistemic}} > \text{T}_{\text{past}} > \text{T}_{\text{future}} > \text{Mood}_{\text{(ir)realis}} > \ldots > \text{Asp}_{\text{perfect}} > \text{Asp}_{\text{progressive}} \ldots$ (Cinque 1999:76)

- The nominalizer *n* is sandwiched between tense and mood.

(22) OV clauses



[4] As OV is obligatorily interpreted as realis, only the past morpheme is shown here.

4 Agents

4.1 Agents as arguments rather than adjuncts

- **Argument**: OV agents serve as arguments in a low A-position.
- **Case**: OV agents bear the genitive case, which is different from the oblique case for adjuncts as used for agents in the passive-like voice and for standards of comparison in comparatives (Shih 2025).

(23) a. Ka damay, kirikiri**=li**.
NOM dish stir.dry**=1SG.GEN**
'That dish is one that I stir-fried.'

b. Ka Muni, ka-dalame**=li**.
NOM Muni STAT-like**=1SG.GEN**
'Muni is one that I like.'

(24) a. Ki-a-ka-dalame **nakwane** ka Muni.
PASS-REAL-STAT-like **1SG.OBL** NOM Muni.
'Muni is liked by me.'

b. Kay Muni, makakipapawlrialringaw **nakwane**.
this Muni more.careful **1SG.OBL**
'Muni is more careful than me.'

- **Obligatoriness**: OV agents cannot be omitted and thus appear as arguments, while agents in the passive-like construction can be omitted and interpreted as existential (Shih 2025).

(25) a. Kirikiri *(ki ina-Ø) ka damay.
stir.fry *(GEN mother-1SG.GEN) NOM dish
'That dish is one that my mother stir-fried.'

b. Ki-a-kirikiri(-nga) (ki ina-Ø) ka damay.
PASS-REAL-stir.fry(-PFV) (OBL mother-1SG.GEN) NOM dish
'The dish has been stir-fried (by my mother).'

- **Word order**: OV agents cannot be freely ordered with themes, suggestive of arguments.

(26) a. Kirikiri [ki ina-Ø] [ka damay].
stir.fry [GEN mother-1SG.GEN] [NOM dish]
'That dish is one that my mother stir-fried.'

b. *Kirikiri [ka damay] [ki ina-Ø].
stir.fry [NOM dish] [GEN mother-1SG.GEN]
'That dish is one that my mother stir-fried.'

- **Cliticization**: AV pronominal agents cliticize to the negation head, but OV pronominal agents do not cliticize to it, showing that OV agents are in a low A-position.

(27) a. Kai=**naku** wa-lrumay ki Lavurase.
NEG=**1SG.NOM** REAL-hit OBL Lavurase
'I didn't hit Lavurase.'

(28) a. Kai ka kirikiri=**su** ka damay.
NEG KA stir.fry=**2SG.GEN** NOM dish
'That dish is not one that you stir-fried.'

b. *Kai=**su** ka kirikiri ka damay.
NEG=**2SG.GEN** KA stir.fry NOM dish
'That dish is not one that you stir-fried.'

4.2 Agents being syntactically inert

- **Argument**: OV agents are syntactically inert and do not participate in A'-movement.[5]
- **Topicalization**: In AV, agents but not themes can be topicalized, whereas in OV, agents cannot be topicalized.

(29) a. **Ka Lavurase**, wa-lrumay ki tawpungu.
NOM Lavurase REAL-hit OBL dog
'Lavurase hit the dog.'

b. *__Ki__ / *__Ka__ **tawpungu**, wa-lrumay ka Lavurase.
*__OBL__ / *__NOM__ **dog** REAL-hit NOM Lavurase
'Lavurase hit the dog.'

(30) a. *__Ki__ / *__Ka__ **ina=Ø**, kirikiri ka damay.
*__GEN__ / *__NOM__ **mother=1SG.GEN** stir.fry NOM dish
'The dish is one that my mother stir-fried.'

- **Relativization**: Agents but not themes can be relativized from AV relative clauses, whereas agents cannot be relativized from OV relative clauses (C.-F. Chen 1999, Sung 2011).

(31) a. M-u<a>bere ka [**wa-lrumay** ki tawpungu ka] lasu.
M-<REAL>escape NOM [**REAL-hit** OBL dog REL] man
'The man that hit the dog escaped.'

b. *M-u<a>bere ka [**wa-lrumay** ka Lavurase ka] tawpungu.
M-<REAL>escape NOM [**REAL-hit** NOM Lavurase REL] dog
'The dog that Lavurase hit escaped.'

(32) a. *Pasawvaladhane ka [**kirikiri** ka damay ka] lasu.
handsome NOM [**stir.fry** NOM dish REL] man
'The man that stir-fried the dish is handsome.'

- ***wh*-formation**: Like most Formosan languages (Aldridge 2002, Tsai 2023), the fronted interrogative is a *wh*-predicate and the rest of the clause is a headless relative clause in Budai Rukai (C.-F. Chen 1999). Agents but not themes can serve as fronted interrogatives with AV headless relative clauses, whereas agents cannot serve as fronted interrogatives with OV headless relative clauses.

(33) a. Aneane [ka **wa-lrumay** ki tawpungu]?
who [NOM **REAL-hit** OBL dog]
'Who hit the dog?'

b. *Manemane [ka **wa-lrumay** ka Lavurase]?
what [NOM **REAL-hit** NOM Lavurase]
'What did Lavurase hit?'

(34) a. *Aneane [ku **kirikiri** ka damay]?
who [NOM **stir.fry** NOM dish]
'Who stir-fried the dish?'

[5] A'-movement in Budai Rukai, like other Austronesian languages, is quite restricted and may behave more like A-movement with [•D•]. For convenience, I still call topicalization, relativization, and *wh*-formation A'-movement.

4.3 Genitive agents/arguments not ergative

- **Argument**: Genitive agents/arguments in OV are not ergative (as assigned to Spec,*v*P or Spec,VoiceP), unlike other Indonesian-type languages (Arka 2003, 2008, Cole and Hermon 2005, Cole et al. 2008, Aldridge 2008, Legate 2014, and a.o.).

- **Genitive arguments in unaccusatives**

- A subset of OV constructions can have result or reason as having syntactic prominence. In this subset, unaccusative predicates occur with a genitive theme (Comp,VP), which suggests genitives in OV are not confined to Spec,*v*P and are thus structural.

(35) a. Kay kela**=li** ki daane=numi, lri-kisalru=aku ku
this come**=1SG.GEN** OBL house=2PL.GEN FUT-borrow=1SG.NOM OBL
paysu.
money
'The reason that I came to your house is to borrow money.'

b. Manemane ku mwadreke **ki** **Lavurase**?
what NOM stumble **GEN Lavurase**
'What reason is it that Lavurase stumbled?' ('What causes Lavurase to stumble?')

- **Causees in causatives**

- In (some) ergative languages, with causands being absolutive, causers and causees are both marked as ergative in causatives. In OV, with causands being nominative-marked, causers are genitive and causees are oblique.

(36) a. Trumai (isolate language in Brazil)
alaweru-**k** hai-**ts** axos disi-ka.
Alaweru-**ERG** 1SG-**ERG** child.ABS hit-CAUS
'Alaweru made me hit the child.' (V. Chen 2017:90 from Guirardello 1999)

b. Budai Rukai
Ka lrulruuy, pa-salri**=li** **iniane** / *=**ini** kibubuli.
NOM box CAUS-move**=1SG.GEN** **3SG.OBL** / *=**3SG.GEN** stealthily
'The box is one that I made him move stealthily.'

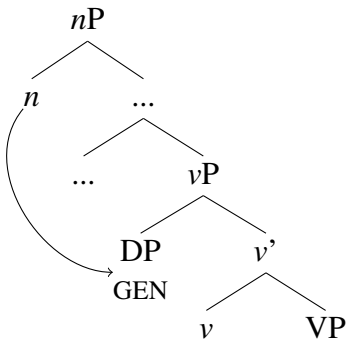
- This suggests that genitives in OV are not ergatives (see also V. Chen 2017 for PV in other Austronesian languages). Note that the embedded domain of causatives is a regular active voice as shown in Appendix A.

4.4 Summary

- Agents are arguments in a low A-position and with syntactically inert behavior but with a structural Case.

- It is thus proposed that OV agents are base-generated arguments in Spec,*v*P within nominalization and receive structural genitive Case from the OV nominalizer.

(37) OV agents



5 Themes

5.1 Theme arguments being syntactically prominent

- **Argument**: OV themes are syntactically prominent, patterning with AV agents in morphosyntactic behavior.
- **Topicalization**: OV themes can be topicalized.

(38) a. **Ka damay**, kirikiri ki ina-Ø.
NOM dish stir.fry GEN mother-1SG.GEN
'That dish is one that my mother stir-fried.'

- **Relativization**: Themes can be relativized from OV relative clauses.

(39) a. Sangu-a-esay ka [**kirikiri** ki ina-Ø ka] damay.
smell-REAL-rich NOM [**stir.fry** GEN mother-1SG.GEN REL] dish
'The dish that my mother stir-fried smells good.'

- ***wh*-formation** Themes can be fronted interrogatives with OV headless relative clauses.

(40) a. Manemane [ka **kirikiri** ki ina-Ø]?
what [NOM **stir.fry** GEN mother-1SG.GEN]
'What did my mother stir-fry?'

- **Case**
- In AV, agents but not themes are marked as nominative; in OV, themes are nominative.

(41) a. Kirikiri ki ina-Ø **ka damay**.
stir.fry GEN mother-1SG.GEN **NOM dish**
'That dish is one that my mother stir-fried.'

b. Ka-dalame ki Lavurase **ka Muni**.
STAT-like GEN Lavurase **NOM Muni**
'Muni is one that Lavurase likes.'

- However, pronominal themes in OV are not realized as nominative but oblique or topic.

(42) a. *Ka-dalam=**aku** ki Lavurase.
STAT-like=**1SG.NOM** GEN Lavurase
'I am one that Lavurase likes.'

b. Ka-dalame ki Lavurase **nakwane**.
STAT-like GEN Lavurase **1SG.OBL**
'I am one that Lavurase likes.'

c. **Kwaku**, ka-dalame ki Lavurase.
1SG.TOP STAT-like GEN Lavurase
'I am one that Lavurase likes.'

- With a higher head, the nominative pronominal theme can surface, but the oblique theme is also grammatical.

(43) a. Kai=**naku** ka ka-dalame ki Lavurase.
NEG=**1SG.NOM** KA STAT-like GEN Lavurase
'I am not one that Lavurase likes.'

b. Kai ka ka-dalame ki Lavurase **nakwane**.
NEG KA STAT-like GEN Lavurase **1SG.OBL**
'I am not one that Lavurase likes.'

5.2 Theme arguments interpreted high

- **Argument**: Themes appear to occupy a higher position than agents, as evidenced by variable binding.
- Main idea: OV themes are interpreted high, (a) if quantificational themes can bind into agents and (b) if themes are not bound into by quantificational agents.
- In AV, quantificational agents can bind into themes, but quantificational themes cannot bind into agents.

(44) a. Ma-dalame [sana ka sinsi*i*] [ki panudhu=ini*i*].
STAT.REAL-like [every KA teacher] [OBL student=3SG.GEN]
'Every teacher*i* likes his*i* students.'

b. Ma-dalame [kay sinsi=ini**i/k*] [sana ka panudhu*i*].
STAT.REAL-like [this teacher=3SG.GEN] [every KA student]
'His**i/k* teacher likes every student*i*.'

- In OV, quantificational themes can bind into agents, whereas pronominal elements inside themes cannot have a co-varying reading with quantificational agents. This suggests that themes occupy a higher position than agents.

(45) a. Ka-dalame [ki t-ina=ini*i*] [sana ka lrulay*i*].
STAT-like [GEN T-mother=3SG.GEN*i*] [every KA child*i*]
'Every child*i* is loved by his*i* mother.'

b. Ka-dalame [sana ka lrulay*i*] [ka t-ina=ini**i/k*].
STAT-like [every KA child] [NOM T-mother=3SG.GEN]
'His**i/k* mother is loved by every child*i*.'

5.3 Theme arguments interpreted low

- **Argument**: Themes appear to occupy a lower position than agents, as evidenced by the Condition C and reflexive binding.
- **Condition C**
- Main idea: OV themes are interpreted low, (a) if agents can induce the Condition C violation to themes and (b) if themes cannot induce the Condition C violation to agents.
- Standard Condition C applies in AV. Note that Budai Rukai has covert third person pronouns, so another R-expression is used for the subject position.[6]

(46) a. Wa-lapu [ku tawpungu=ini*i*] [ka Lavurase*i*].
REAL-keep [OBL dog=3SG.GEN] [NOM Lavurase]
'Lavurase*i* keeps his*i* dog.'

b. Wa-lapu [ku tawpungu ki Lavurase*i*] [ka lasu**i/k*].
REAL-keep [OBL dog GEN Lavurase] [NOM male]
'He/The man**i/k* keeps Lavurase*i*'s dog.'

- In OV, the pronominal agent induces the Condition C violation to themes, whereas the pronominal theme (in oblique, see Section 5.1) does not induce the Condition C violation to agents. This asymmetry suggests that themes are interpreted lower than agents.

[6]This ungrammaticality may be confounded, as two R-expressions in Budai Rukai are less likely to co-refer regardless of c-commanding. Nonetheless, I still regard ungrammaticality here as the Condition C violation.

(47) a. Lapu[=ini*i/k*] [ka tawpungu ki Lavurase*i*].
keep[=3SG.GEN] [NOM dog GEN Lavurase]
'Lavurase*i*'s dog is one that he**i/k* keeps.'

b. Pilri [ki sinsi ki Lavurase*i*] [iniane*i*].
choose [GEN teacher GEN Lavurase] [3SG.OBL]
'He*i* is one that Lavurase*i*'s teacher chose.'

- **Reflexive binding**
- Main idea: OV themes are interpreted low, (a) if the genitive agents can occur with the reflexive theme and (b) if the nominative/oblique themes cannot occur with the reflexive agent.
- Reflexives are expressed by an affix *ngi-*, and may be a semantic Voice^0 that is identified with the argument in Spec,VoiceP, namely AG = TH (see Paparounas 2023 for semantic binding; see Appendix B for a preliminary look of Budai Rukai).
- In AV, agents can occur with reflexive themes.

(48) a. Sa calri-calrig=aku ki angatu, **ngi**-a-pa-pakay=**aku**.
when RED-chop=1SG.NOM OBL tree **REFL**-REAL-RED-hit=**1SG.NOM**
'When I was chopping the tree, I hit myself.'

b. Sa calri-calrigi ka Lavurase ki angatu,
when RED-chop NOM Lavurase OBL tree
ngi-a-pa-pakay=Ø.
REFL-REAL-RED-hit=**3SG.NOM**
'When Lavurase was chopping the tree, he hit himself.'

- In OV, agents can occur with reflexive themes, suggesting that themes have connectivity with the verbal projection within nominalization.

(49) a. Sa calri-calrig=aku ki angatu, **ngi**-pa-pakay=**li**.
when RED-chop=1SG.NOM OBL tree **REFL**-RED-hit=**1SG.GEN**
'When I was chopping the tree, I hit myself.'

b. Sa calri-calrigi ka Lavurase ki angatu, **ngi**-pa-pakay=**ini**.
when RED-chop NOM Lavurase OBL tree **REFL**-RED-hit=**3SG.GEN**
'When Lavurase was chopping the tree, he hit himself.'

- In OV, themes cannot occur with reflexive agents.

(50) a. *Sa calri-calrig=aku ki angatu, **ngi**-pa-pakay=**aku**.
when RED-chop=1SG.NOM OBL tree **REFL**-RED-hit=**1SG.NOM**.
'When I was chopping the tree, I hit myself.'

b. *Sa calri-calrigi ka Lavurase ki angatu, **ngi**-pa-pakay=Ø.
when RED-chop NOM Lavurase OBL tree **REFL**-RED-hit=**3SG.NOM**
'When Lavurase was chopping the tree, he hit himself.'

- Even if themes are realized as nominative on a negation head, they still cannot occur with reflexive agents. This is expected as the PredP is in the matrix clause, and the Pred^0 cannot host the semantic reflexive Voice^0.

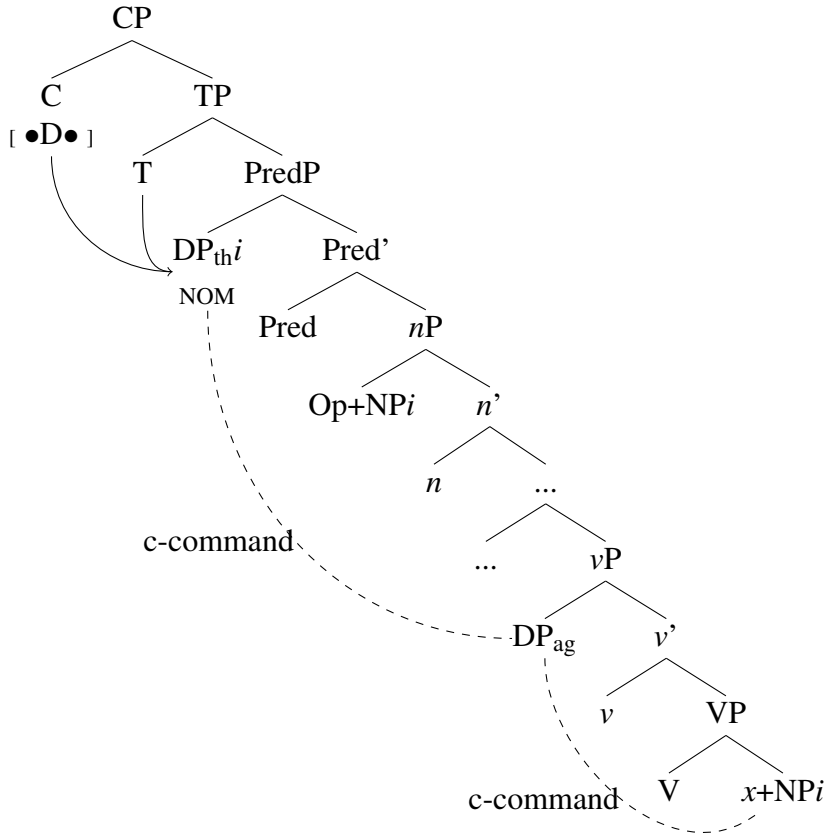
(51) a. Q: Tu-a-tumane kay drapale=su? Ngi-a-dra-dredre=su?
do-REAL-do.what this foot=2SG.GEN REFL-REAL-RED-step=2SG.NOM
'How is your foot? Did you step on yourself?'

b. A: *Ini, ini, kai=**naku** ka **ngi**-dra-dredre.
no no NEG=**1SG.NOM** KA **REFL**-RED-step
'No, I didn't step on myself.'

5.4 Summary

- OV themes can undergo A'-movement and receive the nominative Case, which suggests a higher position that is closer to the corresponding probes.
- OV themes can be interpreted high, as evidenced by variable binding, whereas they can also be interpreted low, as evidenced by the Condition C and reflexive binding.
- It is thus proposed that themes occupy two positions, Spec,PredP and Comp,VP.
  - In Spec,PredP, themes serve as the subject of the nominalized predicate. They receive structural Nominative, are probed by C, and are interpreted high as in variable binding.
  - In Comp,VP, themes are internally complex as a variable taking an NP-complement, which moves to Spec,*n*P as an operator to identify with themes in Spec,PredP. The NP-complement is PF-deleted under the identity with themes in Spec,PredP, but is LF-interpreted without vehicle change (see Salzmann 2019). With NP-complements in Comp,VP, themes are interpreted low for reconstruction effects in the Condition C violation and reflexive binding.

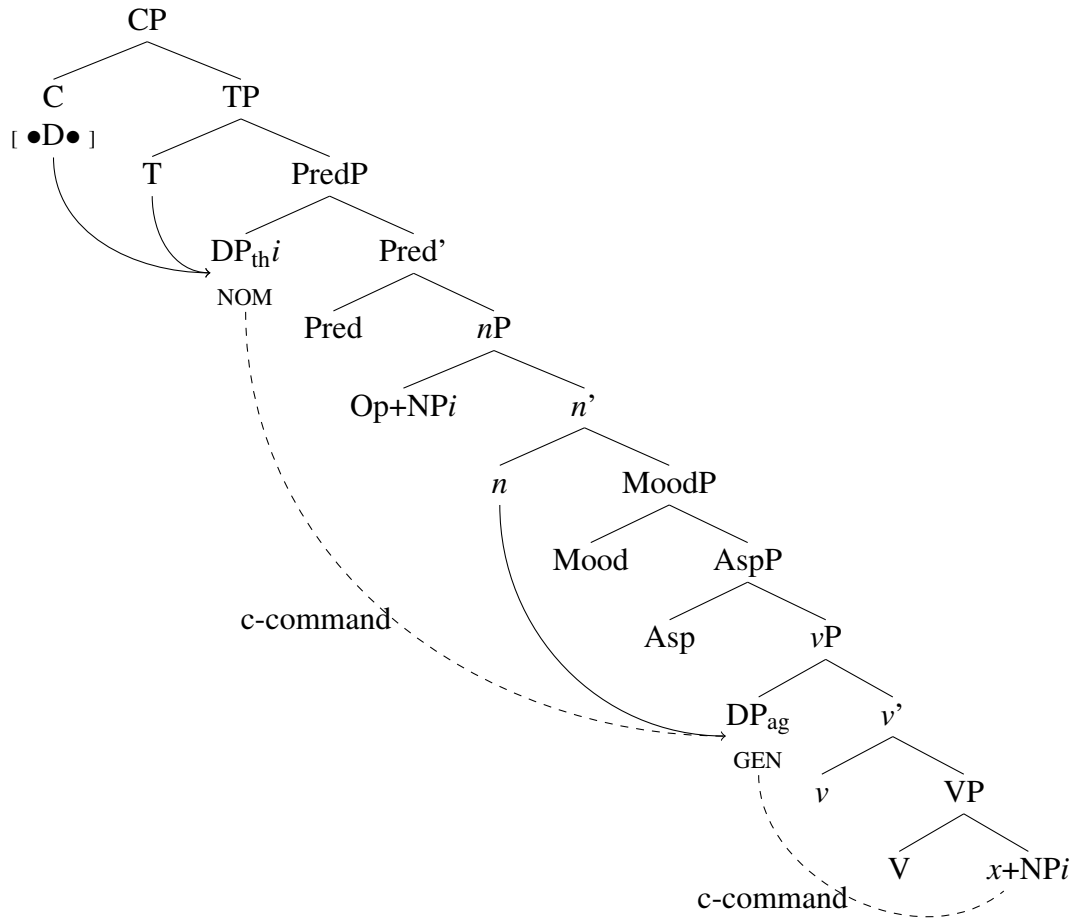
(52) OV themes



6 Analysis

- A biclausal structure is proposed for OV in Budai Rukai. The linearization is set aside here.
- In the matrix clause, the nominalized structure serves as predicates, as in the complement of Pred^0, whereas the theme argument serves as the subject, as base-generated in the Spec,PredP and assigned the nominative Case by T.
- Within nominalization, a full verbal projection is present with AspP and MoodP where agents are assigned the genitive Case by *n*.
- The theme in Comp,VP is internally complex as a(n) operator (Op) / variable (*x*) taking an NP-complement. The NP-complement is PF-deleted but is LF-interpreted for the reconstruction effects (see Salzmann 2019). This theme moves to Spec,*n*P to create λ-abstraction and is filled by the theme in Spec,PredP in semantics.

(53) OV structure in Budai Rukai



- Problem: Why is the Condition C violation not found between the theme in Spec,PredP and the agent, as in (47b) in Section 5.3?

7 Appendix A: Causatives

- The causatives in Budai Rukai are expressed by the morpheme *pa-*, which is shared in many Austronesian languages.
- The embedded domain of causatives in Budai Rukai is an active voice not an applicative voice, as evidenced by agent-oriented adverbs and also instrumental phrases. (See Legate 2014 for the diagnoses, and V. Chen 2017 for Austronesian languages).

(54) a. Pa-a-salri=aku [ki Lavurase / iniane] ka lrulruuy
CAUS-REAL-move=1SG.NOM [OBL Lavurase / 3SG.OBL] OBL box
kibubuli.
stealthily
'I made Lavurase/him move the box stealthily.' (Lavurase did it stealthily.)

b. Pa-a-ridraydr=aku [ki Lavurase / iniane] ka angatu
CAUS-REAL-saw=1SG.NOM [OBL Lavurase / 3SG.OBL] OBL wood
ara-puku.
use-power.
'I made Lavurase/him saw the wood with force.' (Lavurase did it with force.)

(55) a. Pa-a-salri=aku [ki Lavurase / iniane] ka lrulruuy
CAUS-REAL-move=1SG.NOM [OBL Lavurase / 3SG.OBL] OBL box
arakay ku sadrudrudrulrane.
use OBL cart
'I made Lavurase move the box with a cart.' (Lavurase did it with a cart.)

b. Pa-a-ridraydr=aku [ki Lavurase / iniane] ka angatu **arakay**
CAUS-REAL-saw=1SG.NOM [OBL Lavurase / 3SG.OBL] OBL wood **use**
ku ridraydri=li.
OBL saw=1SG.GEN.
'I made Lavurase saw the wood with my saw.' (Lavurase did it with my saw.)

- As a regular verb phrase, causees can thus bind reflexive causands, or causees can occur with a reflexive voice.

(56) a. Ma-salru=aku ki Lavurase.
STAT.REAL-believe=1SG.NOM OBL Lavurase
'I believe in Lavurase.'

b. Ngi-a-ka-ka-salru ka Lavurase.
REFL-REAL-RED-STAT-believe NOM Lavurase
'Lavurase believes in himself.'

c. Pa-ngi-a-ka-ka-salru=aku ki Lavurase.
CAUS-REFL-REAL-RED-STAT-believe=1SG.NOM OBL Lavurase
'I made Lavurase believe in himself.'

8 Appendix B: Reflexives

- As the reflexive in Budai Rukai is expressed as a bound morpheme *ngi-* on the verb, it could be a syntactic argument or could be a semantical Voice^0 (AG = TH).
- The reflexive behavior in Budai Rukai suggests a semantic Voice^0.
- **Proxy reading**: The reflexive *ngi-* does not support the proxy reading.

(57) a. Context: Lavurase saw his picture in a museum, and someone reported:
#Ngi-a-dra-dreele ka Lavurase.
REFL-REAL-RED-see NOM Lavurase
'Lavurase saw himself.'

b. Context: Lavurase saw his statue and took a picture of it, and someone reported:
#Ngi-a-sasiasingi ka Lavurase.
REFL-REAL-RED.photograph NOM Lavurase
'Lavurase photographed himself.'

- **Ditransitives**: The reflexive *ngi-* cannot be recipients.

(58) a. Wa-baay ka Lavurase nakwane ku hungu.
REAL-give NOM Lavurase 1SG.OBL OBL book
'Lavurase gave me a book.'

b. Ngi-a-ba-baay ka Lavurase ku hungu.
REFL-REAL-RED-give NOM Lavurase OBL book
'Lavurase himself gave (someone) a book.'
'#Lavurase gave himself a book.'

- **ECM (causatives)**: The reflexive *ngi-* can serve as causees but not causands.

(59) causee

a. Pa-a-sinav=aku ki Lavurase ku kisi.
CAUS-REAL-wash=1SG.NOM OBL Lavurase OBL bowl
'I made Lavurase do the dishes.'

b. Ngi-a-pa-pa-sinav=aku ku kisi.
REFL-REAL-RED-CAUS-wash=1SG.NOM OBL bowl
'I made myself do the dishes.'

(60) causand

a. Pa-a-lrumadh=aku ki Lavurase ki Laucu.
CAUS-REAL-hit=1SG.NOM OBL Lavurase OBL Laucu
'I made Lavurase hit Laucu.'

b. ??Ngi-a-pa-pa-lrumadh=aku ki Lavurase.
REFL-RED-RED-CAUS-hit=1SG.NOM OBL Lavurase
'I made Lavurase hit myself.'

c. Pa-a-lrumadh=aku ki Lavurase nakwane.
CAUS-REAL-hit=1SG.NOM OBL Lavurase 1SG.OBL
'I made Lavurase hit me.' [my consultant said that she may use this]

References

Aldridge, Edith. (2002). Nominalization and wh-movement in Seediq and Tagalog. *Language and Linguistics*, *3*(2), 393–426.

Aldridge, Edith. (2004). *Ergativity and word order in Austronesian languages* (Doctoral dissertation). Cornell University.

Aldridge, Edith. (2008). Phase-based account of extraction in Indonesian. *Lingua*, *118*(10), 1440–1469.

Aldridge, Edith. (2012). Antipassive and ergativity in Tagalog. *Lingua*, *122*(3), 192–203.

Arka, I. Wayan. (2003). *Balinese morphosyntax: A lexical-functional approach*. Canberra: Pacific Linguistics.

Arka, I. Wayan. (2008). Voice and the Syntax of =a/-a Verbs in Balinese. In P. K. Austin & S. Musgrave (Eds.), *Voice and grammatical relations in austronesian languages* (Vol. 1, pp. 70–89).

Baker, Mark. (1985). The mirror principle and morphosyntactic explanation. *Linguistic inquiry*, *16*(3), 373–415.

Chang, Henry Yungli. (2015). Extractions in Tsou causative applicatives. *Lingua Sinica*, *1*(1), 1–48.

Chang, Melody Ya-yin. (2004). *Subjecthood in Tsou Grammar* (Doctoral dissertation). National Tsing Hua University.

Chen, Cheng-Fu. (1999). *Wh-words as Interrogatives and Indefinites in Rukai* (Master's thesis). National Taiwan University.

Chen, Cheng-Fu. (2008). *Aspect and tense in Rukai: Interpretation and interaction* (Doctoral dissertation). The University of Texas at Austin.

Chen, Victoria. (2017). *A Reexamination of the Philippine-Type Voice System and its Implications for Austronesian Primary-Level Subgrouping* (Doctoral dissertation). University of Hawai'i.

Chen, Victoria and McDonnell, Bradley. (2019). Western Austronesian voice. *Annual Review of Linguistics*, *5*(1), 173–195.

Cinque, Guglielmo. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford University Press.

Cole, Peter, and Hermon, Gabriella. (2005). Subject and non-subject relativization in Indonesian. *Journal of East Asian Linguistics*, *14*(1), 59–88.

Cole, Peter, Hermon, Gabriella, and Yanti. (2008). Voice in Malay/Indonesian. *Lingua*, *118*(10), 1500–1553.

Erlewine, Michael, Levin, Theodore, and van Urk, Coppe. (2017). Ergativity and Austronesian-Type Voice Systems. , 373–396.

Ferrell, Raleigh. (1979). Construction markers and subgrouping of Formosan languages. *Southeast Asian linguistic studies*, *3*, 199–211.

Guo, Qing-Hua. (1979). *Budai Complementation* (Master's thesis). Fu Jen Catholic University.

Himmelmann, Nikolaus. (2001). Voice in western Austronesian: an update. In *The history and typology of western Austronesian voice systems* (pp. 7–16). Pacific Linguistics.

Himmelmann, Nikolaus. (2005). The Austronesian languages of Asia and Madagascar: typological characteristics. In *The Austronesian Languages of Asia and Madagascar* (p. 110-181). Routledge.

Kaufman, Daniel. (2009). Austronesian typology and the nominalist hypothesis. In *Austronesian historical linguistics and culture history* (pp. 187–216). Pacific Linguistics.

Kaufman, Daniel. (2017). Lexical Category and Alignment in Austronesian. In *The Oxford Handbook of Ergativity* (pp. 589–628). Oxford University Press.

Legate, Julie Anne. (2014). *Voice and v: Lessons from acehnese* (Vol. 69). MIT Press.

Nie, Yining. (2020). *Licensing arguments* (Unpublished doctoral dissertation). New York University.

Paparounas, Lefteris. (2023). *Voice from syntax to syncretism* (Doctoral dissertation). University of Pennsylvania.

Patrianto, Hero. (2023). Split to remain symmetrical: A spectrum of Indonesian-type voices. In *Paper presented at Austronesian Formal Linguistics Association 31.*

Pearson, Matthew. (2001). *The clause structure of Malagasy: A minimalist approach* (Unpublished doctoral dissertation). University of California, Los Angeles.

Pearson, Matthew. (2005). The Malagasy subject/topic as an A'-element. *Natural Language & Linguistic Theory*, *23*(2), 381–457.

Rackowski, Andrea. (2002). *The structure of Tagalog: Specificity, voice, and the distribution of arguments* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Rackowski, Andrea and Richards, Norvin. (2005). Phase edge and extraction: A Tagalog case study. *Linguistic inquiry*, *36*(4), 565–599.

Richards, Norvin. (2000). Another look at Tagalog subjects. In (pp. 105–116). Springer.

Salzmann, Martin. (2019). A new version of the Matching Analysis of relative clauses. *Reconstruction effects in relative clauses. Berlin: De Gruyter*, 187–224.

Shih, Chun-Hung. (2022). *Adverbial Constructions in Budai Rukai* (Master's thesis). National Taiwan University.

Shih, Chun-Hung. (2025). Ki-passives in Budai Rukai. *Poster presented at the 101st Linguistics Society of America (LSA 101)*.

Sung, Li-May. (2011). Clausal nominalization in Budai Rukai. *Nominalization in Asian languages: Diachronic and typological perspectives*, 523–559.

Tsai, Wei-Tien Dylan. (2023). Interrogative Clauses in Formosan Languages. In *Handbook of Formosan Languages: The Indigenous Languages of Taiwa.* Brill, online.

Wolff, John U. (1996). The development of the passive verb with pronominal prefix in Western Austronesian languages. In *Reconstruction, classification, description: Festschrift in honor of isidore dyen* (pp. 15–40). Abera Hamburg.

Zeitoun, Elizabeth. (2000). *A Reference Grammar of (Budai) Rukai*. Taipei: Yuangliu.

Zeitoun, Elizabeth. (2023). Rukai. In *Handbook of Formosan Languages: The Indigenous Languages of Taiwan.* Brill, online.