

Learning the organization of the verbal phase in Philippine-type and Indonesian-type languages

Michael Yoshitaka Erlewine, mitcho@nus.edu.sg

National University of Singapore

ISMIL 26, May 2023¹

1 Introduction: Subjecthood on the verbal phase

A long-standing intuition in generative syntax is that structure is built in “chunks.”

- ▶ Chomsky (2000, 2001) introduced the notion of a “*phase*”:

- (1) **Phase-Impenetrability Condition:** (Chomsky, 2000: 108)

For phase head H, “the [complement] of H is not accessible to operations outside HP; only H and its edge [= specifiers and adjuncts] are accessible to such operations.”

- Certain categories are phases, including DPs (noun phrases) and CPs (full clauses)...
- and the *verbal phase*: roughly corresponding to the structure of the verbal predicate, within which all arguments are generated.

Why is this important for the analysis of Austronesian languages?

- Many western Austronesian languages are well known for their “*voice systems*” (see e.g. Erlewine, Levin, and Van Urk 2017, 2020 and Chen and McDonnell 2019):

- (2) **Austronesian-type voice systems:**

- a. Each clause has one privileged argument (“subject”)²;
- b. voice morphology on the verb varies with the choice of subject;
- c. \bar{A} -extraction (*wh*-movement, relativization, etc.) is limited to the subject.

⇒ Extraction facts suggest that the subject must always be in the “edge” of the verbal phase.

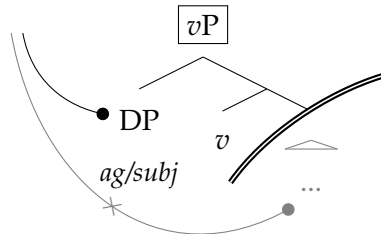
¹ For comments and discussion that informed this work, I especially thank Hiroki Nomoto, Alex Smith, Carly Sommerlot. This work is supported by the Singapore Ministry of Education and the National University of Singapore under grants A-8000132-00-00 and A-8001136-00-00. Errors are mine.

² Often also called the “pivot.”

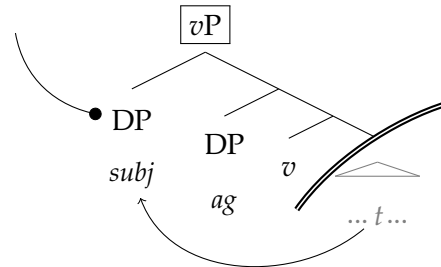
In much phase-based work including Chomsky (2000, 2001), a single head — “little *v*” — introduces the external argument (agent/actor) and also serves as the phase edge, allowing lower arguments to be accessible to operations from above:

(3) **The verbal phase with one head: *v***

a. Actor Voice (AV):



b. Non-Actor Voices (NAV):



- Precisely this geometry for the verbal phase has been adopted in Minimalist analyses for Philippine-type languages (e.g. Aldridge, 2004; Rackowski and Richards, 2005; Erlewine and Levin, 2021; Erlewine and Lim, to appear) and also for Indonesian-type languages (e.g. Aldridge, 2008; Sato, 2012; Erlewine et al., 2020).

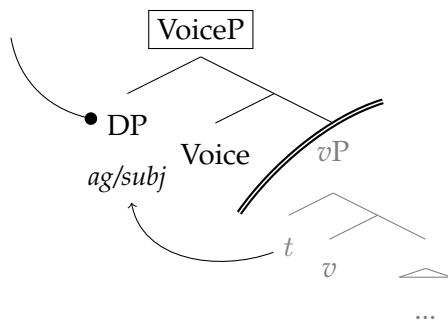
An alternative approach would be to ascribe these two functions to two different heads:

- *v* introduces the agent;
- Voice (above *v*) serves as the phase head.

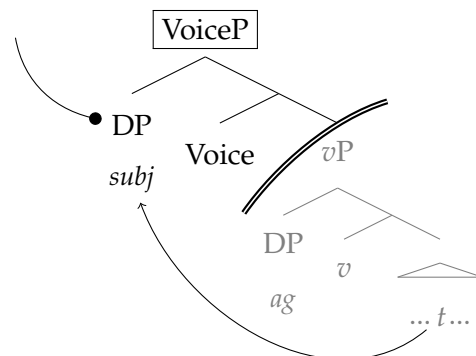
(See e.g. Collins 2005; Gallego 2008; Richards 2010; Coon, Mateo Pedro, and Preminger 2014.)

(4) **The verbal phase with two heads: Voice and *v***

a. Actor Voice (AV):



b. Non-Actor Voices (NAV):



Specifically, suppose that Voice always hosts exactly one nominal specifier: the subject.

Today

I discuss two major subtypes of western Austronesian languages (see e.g. Wolff, 1996; Himmelmann, 2005; Blust, 2013; Chen and McDonnell, 2019):

- **Philippine-type languages instantiate the one-head theory (3), whereas Indonesian-type languages instantiate the two-head theory (4).**

- A range of different morphological and syntactic features of these languages lead learners to these organizations of the verbal phase.

2 Philippine-type languages

2.1 Evidence from clitics

Philippine-type languages commonly have **second-position clitics**:

(5) **Second-position clitic pronouns in Tagalog:**

- | | | | | | | | | |
|----|---------------------|---------|---------|----|----------------------|---------|---------|---------------|
| a. | Tu~tulung-an | =mo | =ako. | b. | Hindi | =mo | =ako | tu~tulung-an. |
| | FUT-help-LV | GEN.2SG | NOM.1SG | | NEG | GEN.2SG | NOM.1SG | FUT-help-LV |
| | 'You will help me.' | | | | 'You won't help me.' | | | |

As discussed in Erlewine and Levin 2021, the following holds in all Philippine-type languages:

(6) **Generalization:** (Erlewine and Levin, 2021: 412)

"In transitive clauses, second-position clitic pronouns in Philippine-type languages are limited to pivot [= subject] arguments and non-pivot [= non-subject] agents."

This generalization is not explained simply by proposing that pronouns in certain cases (e.g. "accusative" for non-subject themes) lack clitic forms. (Henrison Hsieh, p.c. in *ibid.*: 411–412.)

(7) **Demonstrative pronouns can be clitics or non-clitics:**

- | | | | | | | |
|-----|--|-----|-------|-----------------------|------------|------------------------------|
| Ang | lalaki | ang | hindi | { [✓] =nito} | na-kita | { [✓] nito}. |
| NOM | boy | NOM | NEG | GEN.PROX | PV.PFV-SEE | GEN.PROX |
| | 'It's the boy that this one didn't see.' | | | | | (genitive non-subject agent) |

(8) **Non-subject theme genitive must be a non-clitic:**

- | | | | | | | |
|-----|--|-----|-------|----------|------------|------------------------------|
| Ang | lalaki | ang | hindi | {*=nito} | naka-kita | { [✓] nito}. |
| NOM | boy | NOM | NEG | GEN.PROX | AV.PFV-SEE | GEN.PROX |
| | 'It's the boy that didn't see this one.' | | | | | (genitive non-subject theme) |

- ▶ Erlewine and Levin 2021: The generalization in (6) is immediately explained by the one-head theory of the verbal phase (3).
 - Clitic pronouns can clearly be outside of the verbal phase (e.g. hosted on negation).
 - Assuming that clitic pronouns reflect the movement of (or agreement with) pronouns in argument positions, clitic pronouns are only possible for arguments at the edge of the verbal phase: only the agent subject in Actor Voice (3a), and only the subject and non-subject agent in Non-Actor Voices (3b).
- In turn, the existence of clitic pronouns with this distribution provides direct evidence for the learner for the one-head theory of the verbal phase as in (3).

2.2 Evidence from extraction

- \bar{A} -extraction (*wh*-movement, relativization, topicalization, etc.) in many Philippine-type languages has been traditionally described as limited to the subject.
- ▶ However, certain \bar{A} -constructions in certain languages also **allow the movement of non-subject agents**, as well as subjects.

Topicalization of non-subject agents is in fact well attested across many Philippine languages, whereas topicalization of non-subject themes is ungrammatical or unattested. (Topics in such constructions appear in nominative, despite corresponding to a postverbal genitive position.)

(9) Tagalog (De Guzman, 1995: 57):³

- a. **Ang nanay**, lulutu-in (=niya) ang isda sa kusina.
 NOM mother cook-PV GEN.3sg NOM fish DAT kitchen
 ‘The mother, (she) will cook the fish in the kitche.’
- b. ***Ang isda**, mag-lulutu ang nanay sa kusina.
 NOM fish AV-cook NOM mother
 Intended: ‘The fish, mother will cook (it).’

- See also similar examples in Hiligaynon (Mithun, 2019: 159), Limos Kalinga (Ferreirinho, 1993: 68–71), Kapampangan (Mirikitani, 1972: 154; Rowsell, 1983: 57–58), Pangasinan (Benton, 1971: 154), and Western Subanon (Blake, 2020). Reid (1978: 36) also presents parallel examples of this form from Bontok, Ilokano, Ivatan, and Tagalog.
- ▶ See especially Erlewine and Lim to appear, which establishes that **topicalization involves movement in Bikol**, targeting subjects and non-subject agents but not non-subject themes.

³ Latrouite (2011: 69) reports a variant of example (9b) with a post-verbal genitive demonstrative pronoun *nito* with the judgment mark ???. Richards (2005) observes the same contrast with *ay*-topicalization.

Relativization of non-subject agents is also attested, counter to conventional wisdom (the subject-only \bar{A} -extraction restriction).

(10) **Relativization from Tagalog Locative Voice clause:** (Hsieh, to appear)

- a. ✓ damit=ng [ni-labh-an ng bata ___NOM]
 clothes=LK PFV-laundry-LV GEN child
 ‘clothes that the child washed’
- b. ?bata=ng [ni-labh-an ___GEN ang damit]
 child=LK PFV-laundry-LV NOM clothes
 ‘child who washed the clothes’

- For Tagalog, see Tanaka et al. (2016), Pizarro-Guevara and Wagers (2018), and Hsieh (2020, to appear). Bondoc (2020) reports that non-subject agent relativization is also possible in Akeanon and Cebuano, but not non-pivot theme relativization.

2.3 Summary

- Non-subject agents in Philippine(-type) languages participate in a range of constructions that involve their movement (or agreement) out of the verbal phase:
 - second-position clitic pronouns (extremely common)
 - topicalization (well attested)
 - relativization (marked / only in some languages)
- ▶ **Non-subject agents should be in the edge of the verbal phase**, but not non-subject themes \Rightarrow the one-head theory (3).

3 Indonesian-type languages

- ▶ I propose that (many) **Indonesian-type languages utilize the two-head theory** for the verbal phase, as in (4).
 - Voice \leftrightarrow *me-*, *di-*, (*ber-*, *ter-*, ...)
 - *v* \leftrightarrow *N-*
- Central to this hypothesis is the idea of **decomposing *meN-***, following Gil 2002, Benjamin 2009, and citations there. Again, I hypothesize that the morphology and syntax of Indonesian-type languages **provide evidence that guides learners** to the two-head theory.
- To my knowledge, no prior work has explored the two-head theory for Indonesian-type languages except Sommerlot 2020.

3.1 Evidence from morphology

Nominalizations in Standard Malay and Standard Indonesian (SM/SI) may involve *peN-* or *pe-*, largely corresponding to verbal predicates with *meN-* vs *ber-*, respectively. (See e.g. Denistia 2020 and citations there.)

- (11) a. *mengajar* 'teach' ~ *pengajar* 'teacher' (Nomoto, 2017)
b. *belajar* 'learn' ~ *pelajar* 'student'
- (12) a. *pengasih* 'one who is loving' (Hassan 1974, in Benjamin 2009: 304)
b. *pekasih* 'one who is loved'

- Such correspondences support parsing *N-* as a shared agent-related morpheme in both *meN-* and *peN-*. (See Benjamin 2009: 303–304 for a suggestion along these lines.)

Verbal reduplication of active verbs in SM/SI includes the nasal *N-* when it undergoes coalescence with stem-initial consonant, but does not include *me-* (see e.g. Lapoliwa 1981).

- (13) active *menulis* 'write' > active *menulis-nulis* 'write repeatedly' (SM/SI)
(cf **menulis-tulis*, **menulis-menulis*)

- As Benjamin (2009: 298) notes (crediting Hendon (1966: 46–47) for the idea), the analysis of such reduplication is simplified if such forms are actually an active prefix *me-* on a reduplicated stem *nulis-nulis*, prenasalized in the context of active voice.

The analysis above is complicated by the fact that, for stems where coalescence does not occur, both *N* and the stem-initial consonant remain, with reduplication applying only to the stem:

- (14) *baca* 'read' > active *membaca-baca* 'read repeatedly' (SM/SI)
(cf **membaca-mbaca*)

However, these facts are different in many regional and colloquial Malay/Indonesian varieties, often inviting more straightforward parses between the two heads:

- In Riau Indonesian (Gil, 2002: 258–259), the “active prefix” may be *me-* alone or *N-* alone, phonologically conditioned. *N-* is included in reduplication but *me-* is not:

- (15) a. *minjam* 'borrow' > *minjam-minjam* 'borrow repeatedly' (cf **minjam-pinjam*)
b. *lempar* 'throw' > *melempar-lempar* 'throw repeatedly' (cf **melempar-melempar*)

- In colloquial Johor Malay (Onn, 1976: 178), where coalescence does not occur, the *N* part is optionally included in reduplication:

- (16) a. *gali* 'dig' > *menggali-(ng)gali* 'dig continuously'
 b. *tari* 'dance' > *menari-nari* 'dance repeatedly'

This reflects the SM/SI analysis above, but with optionality in *N-* as part of Voice or *v* where coalescence does not occur.

Association of *me-* and *N-* with distinct functions is evident in various regional and colloquial varieties of Malay/Indonesian and other Malayic languages, as discussed in Gil 2002, Benjamin 2009, and others. This even leads to the possibility in some varieties of *di-N-V* forms:

- (17) a. *potong* 'cut' > *di-motong-nya* (Riau Indonesian; Gil 2002: 265)
 b. *pinjam* 'borrow' > *di-minjam*
- (18) a. *bunuh* 'kill' > *di-munuh* (Salako Kendayan (Malayic; W. Borneo))
 b. *rumput* 'weed' > *di-nga-rumput* (Adelaar, 2005: 218–219)

- Such forms directly motivate the segmentation of *meN-* into *me-* and *N-*, with *me-* occupying the same position as *di-*.

3.2 Evidence from the position of agents

The position of agents in Indonesian-type languages also supports the two-head theory:

- To my knowledge, \bar{A} -extraction (e.g. relativization, clefting) non-subject agents is never allowed in any Indonesian-type language, unlike in the Philippine-type languages above.
- On the other hand, non-subject agents can appear before the verb, i.e. in the "bare passive."
- How can agents be preverbal but still inaccessible for \bar{A} -extraction?
 ⇒ Agents must have a position before (above) the verb but not at the phase edge, suggesting the two-head theory.

Low agents (below auxiliaries, immediately preverbal), as in the bare passive, generally cannot cooccur with any voice prefix. But there are some exceptions:

- Suak Mansi Desa (Malayic, West Kalimantan; Sommerlot 2020) active verbs may appear with *meN-* or *N-* in free variation. However:

(19) **Only *N-* is possible in object extractions:**

Buku to yang opa'-ku {boli / moli / *memoli}.
 book DEM C father-1sg buy N-buy meN-buy
 'This is the book that my father bought.'

(20) **Agents are low in object extractions:**

Opai yang {*inya} **nda'** {inya} milau?
 what C NEG 3sg N-look.for
 'What isn't s/he looking for?'

- Salako Kendayan (Malayic, West Borneo; Adelaar 2005) and Matéq (Land Dayak/Bidayuh, West Kalimantan; Connell 2013) allow for "*di/ni* agent (*N-*)V" patterns:

(21) **Salako Kendayan "*di* agent *N-V-RED*"**: (Adelaar, 2005: 218)

Aŋkoà-lah tuàkŋ kalen di=kau matàh-matàh aŋkoà.
 DIST-EMPH bone catfish DI=2sg N-break-RED DIST
 'That's the catfish-bone you've broken into many pieces.'

(22) **Two *ni* passives in Matéq**: (Connell, 2013: 118)

a. balo rua karék ni *koq* mpulua matéq-éh.
 QUANT seed rubber NI 1sg gather just.before
 'I gathered some rubber seeds earlier.'

b. balo rua karék ni-mpulua *koq* matéq.
 QUANT seed rubber NI-gather 1sg soon
 'I'll gather some rubber seeds later.'

- Such evidence (where available) further motivate an organization where low agents follow *me/di-* and precede *N-*, not at the edge of the phase:

	(<i>me-/di-/...</i>)	agent	(<i>N-</i>)V	
[(phase edge)	Voice	[<i>v+V</i>	...

3.3 Summary

- A range of morphological evidence lead learners of Malay(ic) languages to **decompose** *meN-* as *me-* + *N-* (as per Gil, 2002; Benjamin, 2009).
- Combined with the complete absence of non-subject agent movement and the position of low agents, these facts lead learners to posit the two-head model of the verbal phase.

4 Conclusion

- Austronesian languages are known for their “voice systems,” including the restriction of \bar{A} -movement to the subject.
- The *verbal phase* offers a framework for understanding what makes subjects special.
- ▶ Features of Philippine-type versus Indonesian-type grammars may be cues for the learner to respectively adopt the one-head theory versus two-head theory for the verbal phase.

Terima kasih!

References

- Adelaar, Alexander. 2005. Structural diversity in the Malayic subgroup. In Adelaar and Himmelmann (2005), 202–226.
- Adelaar, Alexander, and Nikolaus P. Himmelmann, ed. 2005. *The Austronesian languages of Asia and Madagascar*. Routledge.
- 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* 118:1440–1469.
- Benjamin, Geoffrey. 2009. Affixes, Austronesian and iconicity in Malay. *Bijdragen tot de Taal-, Land- en Volkenkunde* 165:291–323.
- Benton, Richard A. 1971. *Pangasinan reference grammar*. University of Hawai'i Press. URL <https://hdl.handle.net/10125/62898>, retypeset in 2019.
- Blake, A. L. 2020. Left-dislocation in Western Subanon. *Philippine Journal of Linguistics* 51:30–54.
- Blust, Robert. 2013. *The Austronesian languages*. Asia-Pacific Linguistics, revised edition.
- Bondoc, Ivan Paul M. 2020. Relativization asymmetries in Philippine-type languages: A preliminary investigation. *The Archive: A Journal Dedicated to the Study of Philippine Languages and Dialects* 1:1–34.

- Chen, Victoria, and Bradley McDonnell. 2019. Western Austronesian voice. *Annual Review of Linguistics* 5:173–195.
- Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. MIT Press.
- Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8:81–120.
- Connell, Timothy M. 2013. A sketch grammar of Matéq: A Land Dayak language of West Kalimantan, Indonesia. Master's thesis, University of Canterbury.
- Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. 2014. The role of case in A-bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation* 14:179–242.
- De Guzman, Videia P. 1995. Experiencer verbs in Tagalog. In *Grammatical relations: Theoretical approaches to empirical questions*, ed. Clifford S. Burgess, Katarzyna Dziwirek, and Donna B. Gerds, 45–62. CSLI Publications.
- Denistia, Karlina. 2020. Quantitative studies on the Indonesian prefixes *pe-* and *pen-*. Doctoral Dissertation, Universität Tübingen.
- Erlewine, Michael Yoshitaka, and Theodore Levin. 2021. Philippine clitic pronouns and the lower phase edge. *Linguistic Inquiry* 52:408–425.
- 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*, ed. Ileana Paul, 71–87.
- Erlewine, Michael Yoshitaka, and Cheryl Lim. to appear. Bikol clefts and topics and the Austronesian extraction restriction. *Natural Language & Linguistic Theory* .
- Ferreirinho, Naomi. 1993. *Selected topics in the grammar of Limos Kalinga, the Philippines*. Pacific Linguistics.
- Gallego, Ángel. 2008. Four reasons to push down the external argument. Manuscript, Universitat Autònoma de Barcelona.
- Gil, David. 2002. The prefixes *di-* and *n-* in Malay/Indonesian dialects. In *The history and typology of western Austronesian voice systems*, ed. Fay Wouk and Malcolm Ross, 241–283. Canberra: Pacific Linguistics.
- Hassan, Abdullah. 1974. *The morphology of Malay*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

- Hendon, Rufus S. 1966. *The phonology and morphology of Ulu Muar Malay (Kuala Pilah District, Negri Sembilan, Malaya)*. Number 70 in Yale University Publications in Anthropology. Department of Anthropology, Yale University.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Adelaar and Himmelmann (2005), 110–181.
- Hsieh, Henrison. 2020. Beyond nominative: A broader view of \bar{A} -dependencies in Tagalog. Doctoral Dissertation, McGill University.
- Hsieh, Henrison. to appear. Locality in exceptional Tagalog A' -extraction. *Linguistic Inquiry*.
- Lapoliwa, Hans. 1981. A generative approach to the phonology of bahasa indonesia. Doctoral Dissertation.
- Latrouite, Anja. 2011. Voice and case in Tagalog: The coding of prominence and orientation. Doctoral Dissertation, Heinrich-Heine Universität.
- Mirikitani, Leatrice T. 1972. *Kapampangan syntax*. University of Hawaii Press.
- Mithun, Marianne. 2019. Grammatical relations in Hiligaynon. In *Argument selectors: A new perspective on grammatical relations*, ed. Alena Witzlack-Makarevich and Balthazar Bickel, 131–184. John Benjamins.
- Nomoto, Hiroki. 2017. Sintaksis nominalisasi bahasa Melayu [the syntax of Malay nominalization]. In *Aspek teori sintaksis bahasa melayu*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Onn, Farid. 1976. Aspects of Malay phonology and morphology: a generative approach. Doctoral Dissertation, University of Illinois at Urbana-Champaign.
- Pizarro-Guevara, Jed, and Matt Wagers. 2018. Agent extraction under patient voice in Tagalog is acceptable: Evidence from acceptability ratings in various A' -dependencies. Presented at AFLA 25.
- Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565–599.
- Reid, Lawrence A. 1978. Problems in the reconstruction of Proto-Philippine construction markers. In *Proceedings of ICAL 2*, volume 1, 33–66. Pacific Linguistics.
- Richards, Norvin. 2005. Person-case effects in Tagalog and the nature of long-distance extraction. In *Proceedings of AFLA 12*, ed. Jeffrey Heinz and Dimitrios Ntelitheos, 383–394.
- Richards, Norvin. 2010. *Uttering trees*. MIT Press.
- Rowell, Lorna V. 1983. An ergative analysis of Kapampangan. Master's thesis, University of Calgary.
- Sato, Yosuke. 2012. Successive cyclicity at the syntax-morphology interface: Evidence from Standard Indonesian and Kendal Javanese. *Studia Linguistica* 66:32–57.
- Sommerlot, Carly J. 2020. A reanalysis of the Austronesian nasal prefix: Evidence from Desa,

a Malayic language of West Kalimantan. Presented at WCCFL 38.

- Tanaka, Nozomi, William O'Grady, Kamil Deen, Chae-Eun Kim, Ryoko Hattori, Ivan Paul M. Bondoc, and Jennifer U. Soriano. 2016. An agent advantage in Tagalog relative clause comprehension. In *Proceedings of AFLA 22*, ed. Henrison Hsieh, 191–201.
- 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*, ed. Bernd Nothofer, 15–40. Abera.