

# What makes a voice system?

## On the relationship between voice marking and case\*

Michael Yoshitaka Erlewine, Theodore Levin, Coppe van Urk; MIT

{mitcho,tedlevin,coppe}@mit.edu

Austronesian Formal Linguistics Association 21, University of Hawai'i, May 2014

### 1 Introduction: extraction marking and voice systems

Many languages morphologically mark the difference between **non-subject** (a) and **subject** (b) extraction:

- (1) **English T-to-C movement:**
  - a. Who **did** Alex see?
  - b. Who saw Alex?
- (2) **French *quel/qui* alternation:**
  - a. Qui penses-tu [**que** Marie a rencontré]?  
who think-you that Marie has met  
'Who do you think Marie has met?'
  - b. Qui penses-tu [**qui** a rencontré Marie]?  
who think-you that has met Marie  
'Who do you think has met Marie?'
- (3) **Agent Focus in Kaqchikel and other Mayan languages (Erlewine, 2014):**
  - a. Achike xutěj ri a Juan?  
what ate the CL Juan  
'What did Juan eat?'
  - b. Achike xtj-ǝ ri wäy?  
who ate-AF the tortilla  
'Who ate the tortilla?'
- (4) **Moro *wh-concord* (Rohde, 2006; Rose et al., 2014):**
  - a. ɲwədʒeki (nə).Kuku (nə).gətaðoŋo?  
who (WH).Kuku (WH).abandon  
'Who did Kuku abandon?'
  - b. ɲwədʒeki gətaðo Kuku?  
who abandon K.  
'Who abandoned Kuku?'

☞ We refer to such morphology above as "**extraction marking**."

Many Austronesian languages have a **more articulated** form of extraction marking:

- (5) **Tagalog voice marking (adapted from Guilfoyle et al. 1992):**
  - a. Sino ang b<um>ili ng damit para sa bata?  
who COMP AV.ASP-buy ACC dress for OBL child  
'Who bought the dress for the child?' *Actor Voice (AV)*
  - b. Ano ang b<in>ili ng tao para sa bata?  
what COMP PV.ASP-buy GEN man for OBL girl  
'What did the man buy for the girl?' *Patient Voice (PV)*
  - c. Sino ang i-b<in>ili ng tao ng damit?  
what COMP BV.ASP-buy GEN man ACC dress  
'Who was bought the dress (for) by the man?' *Benefactive Voice (BV)*

☞ We refer to the argument cross-referenced by voice morphology as "**subject**," and refer to movement to this subject position as "**extraction**."

Two common approaches to Austronesian voice systems:

1. Voice morphology **marks** extraction, via (*wh*-)agreement (e.g. Chung, 1994; Richards, 2000; Pearson, 2001, 2005);
2. Voice morphology **feeds** extraction, by means of argument structure alternations (e.g. Guilfoyle et al., 1992; Aldridge, 2004, 2008; Legate, 2012).

The two positions are difficult to distinguish, because of the prominent **one-to-one correspondence** of voice, case, and extraction.

### Today

- We present two systems in which this one-to-one correspondence **breaks down**, in the Nilotic language Dinka and Indonesian-type languages like Balinese.
  - In Dinka, we can dissociate voice and case.
  - In Balinese, we can dissociate voice and extraction.
- We argue that voice system morphology functions as **extraction marking**, just as in the languages in (1–4).
- We propose that voice affects case because **the "subject" position** (the argument referenced by voice morphology) is a **mixed A/ $\bar{A}$ -position**.
  - This position licenses the actor in Actor Voice.
  - In Non-Actor Voices, **the actor lacks a licenser**. Voice system languages *vary* in how they respond to this problem. We'll present examples of **three different strategies**.

\*We thank Edith Aldridge, Julie Legate, David Pesetsky, Norvin Richards for helpful comments. All errors are ours.

## 2 Dissociating voice and case in Dinka

If voice morphology is extraction marking, we expect **dissociations between voice and case**, since voice would not directly determine case.

☞ Here we turn to **Dinka**, a Nilotic language of South Sudan, with a voice system highly reminiscent of Austronesian (Van Urk & Richards, to appear).

### The Dinka voice system

Dinka is a V2 language. It has three voices, which reflect the grammatical function of the **clause-initial “subject” position**:

- (6) a. Ayén a-cé cuín cà̄m ne pà̄l.  
Ayen.ABS 3S-PRF.AV food.ABS eat.NF P knife.ABS  
‘Ayen has eaten food with a knife.’ *Actor Voice (AV)*
- b. Cuín a-cíi Áyèn cà̄m ne pà̄l.  
food.ABS 3S-PRF.PV Ayen.NOM eat.NF P knife.ABS  
‘Food, Ayen has eaten with a knife.’ *Patient Voice (PV)*
- c. Pà̄l a-cénné Áyèn cuín cà̄m.  
knife.ABS 3S-PRF.OBLV Ayen.NOM food.ABS eat.NF  
‘With a knife, Ayen has eaten food.’ *Oblique Voice (OblV)*

The clause-initial subject appears in the unmarked case, usually called “absolutive” in the Nilotic literature (e.g. Dimmendaal, 1985; Andersen, 1991, 2002).

As in many Austronesian languages, voice marking restricts overt  $\bar{A}$ -extraction. Only the subject can be  $\bar{A}$ -extracted:

- (7) a. Yeṇà cé cuín cà̄m ne pà̄l?  
who PRF.AV food.ABS eat.NF P knife.ABS  
‘Who has eaten food with a knife?’
- b. Yeṇú cíi Áyèn cà̄m ne pà̄l.  
what PRF.PV Ayen.NOM eat.NF P knife.ABS  
‘What has Ayen eaten with a knife?’
- c. Yeṇú cénné Áyèn cuín cà̄m.  
what PRF.OBLV Ayen.NOM food.ABS eat.NF  
‘What has Ayen eaten food with?’

Non-initial actors appear in a dedicated case, the “marked nominative” (we’ll get back to this), while non-initial patients are unmarked (i.e. absolutive).

### Voice is independent of case

There are several syntactic environments in which clauses **do not front any constituent to the initial “subject” position**, resulting in a V/Aux-initial clause.

☞ In such cases, **default voice marking** (Actor Voice) is used, but nominals are **case-marked** as they would be in a NAV clause.

- (8) Yes-no questions, which are V1 as in many V2 languages:
- a. **Nhià̄r** Máyèn Adít?  
**love.AV** Mayen.NOM Adit.ABS  
‘Does Mayen love Adit?’
- b. Adít a-gḡi [ná **nhià̄r** Máyèn yè̄n].  
Adit 3S-wonder.AV whether **love.AV** Mayen.NOM 3S.ABS  
‘Adit wonders whether Mayen loves her.’
- (9) Complement clauses with the complementizer *ye*:  
A-cá luéel, [ye **nhià̄r** Cán wít].  
3S-PRF.1SG say c **love.AV** Can.NOM wrestling.ABS  
‘I said that Can loves wrestling.’
- (10) Non-finite clauses, which are headed by the future auxiliary *bé*:  
Bòl a-kó̄r [**bé** Máyèn akokóol gḡ̄r].  
Bol.ABS 3S-want.AV **FUT.AV** Mayen.NOM story.ABS write.NF  
‘Bol wants Mayen to write a story.’

In these cases, moving the actor to clause-initial position is ungrammatical (cf. 10):

- (11) \*Bòl a-cé Ayén lḡ̄ [Mayén (a-)bé akokóol gḡ̄r].  
Bol.ABS 3S-PRF.AV Ayen.ABS beg.NF Mayen.ABS (3S-)FUT.AV story.ABS write.NF  
Intended: ‘Bol has begged Ayen for Mayen to write a story.’

☞ Extraction out of a non-finite clause (e.g. 10 above) triggers voice morphology, **even though such clauses otherwise lack a subject position**:

- (12) a. Yeṇú lḡ̄n-kú Ayén [**bíi** Máyèn gḡ̄r]?  
what beg-1PL Ayen.ABS **FUT.PV** Mayen.NOM write  
‘What are we begging Ayen for Mayen to write?’
- b. Yeṇà lḡ̄n-kú Ayén [**bé** akokóol gḡ̄r]?  
who beg-1PL Ayen.ABS **FUT.AV** story.ABS write.NF  
‘Who are we begging Ayen for \_\_ to write a story?’

**Conclusion:** Case marking is *independent of the voice system*, as they can be dissociated in clauses without an initial “subject” position.

### 3 Voice and multiple extraction in Balinese

We now present a **dissociation between voice and extraction** in Indonesian-type languages, such as Bahasa Indonesia (Chung, 1976; Cole & Hermon, 2005), Jambi Malay (Yanti, 2010), and Balinese.

We focus on **Balinese**. Balinese has two active voices:

- (13) a. Polisi **ng**-ejuk Nyoman.  
 police **AV**-arrest Nyoman  
 'A policeman arrested Nyoman.' *Actor Voice (AV)*
- b. Nyoman **Ø**-ejuk polisi.  
 Nyoman **PV**-arrest police  
 'A policeman arrested Nyoman.' *Patient Voice (PV)*

Voice imposes an extraction restriction:

- (14) Actor extraction ⇒ AV:  
 Nyen **ng**/\***Ø**-alih ci ditu ibi?  
 who **AV**/\***PV**-search you there yesterday  
 'Who looked for you there yesterday?'

- (15) Patient extraction ⇒ PV:  
 Apa \***ng**/\***Ø**-alih ci ditu ibi?  
 what \***AV**/**PV**-search you there yesterday  
 'What did you search for there yesterday?'

**However:** *Multiple extraction is also possible!* In such cases, the verb is PV:

- (16) Buku cen Nyoman \***N**/\***Ø**-paca?  
 book which Nyoman \***AV**/**PV**-read  
 'Which book did Nyoman read?'

The actor must be the immediately preverbal argument in multiple extractions.

- (17) \*Nyen buku ne **N**/\***Ø**-paca?  
 who book DEF **AV**/**PV**-read  
 Intended: 'Who read the book?'

**Conclusion:** We can explain this if we view voice as extraction marking. *Patient voice surfaces whenever a patient is extracted.*

### 4 The relationship between voice and case

☞ The dissociations above support the view that **voice marking is  $\bar{A}$ -extraction marking** in voice system languages (Chung, 1994; Richards, 2000; Pearson, 2001, 2005).

- However, unlike extraction marking in non-voice system languages (English, French, Kaqchikel, Moro in §1), voice often has **repercussions for case**.

**Proposal:** In languages with voice systems, one argument moves to the “subject” position, which is a **mixed  $\bar{A}/\bar{A}$ -position**. Here we will call this subject position Spec,CP.

- Internal arguments are licensed in-situ. The external argument actor is not.
- In AV, the actor is licensed in the subject position.
- In NAV, the external argument actor must be licensed in some other way.

The problem of licensing the external argument actor when it is not moved to the subject position is handled differently in different languages. We demonstrate **three strategies** for licensing the actor in NAV:

1. Ergativity in Formosan/Philippine-type voice systems
2. Oblique (prepositional) case in the Dinka voice system
3. Pseudo-noun incorporation in the Balinese voice system

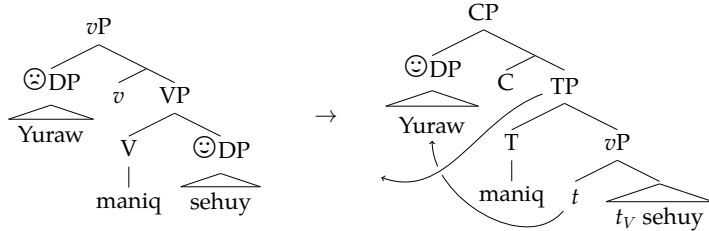
Even genetically close languages may use different strategies, while some genetically distant languages (e.g. Formosan/Philippine-type and Dinka) use conceptually very similar strategies.

## 5 Strategy 1: ergativity

We begin by demonstrating this system with **Atayal** (Formosan; Taiwan).

### Actor Voice derivation

- (18) M-anìq sehuy (qu) Yuraw.  
 AV-eat taro QU Yuraw  
 'Yuraw eats taro.'



TP-fronting yields the observed word order (Aldridge, 2004). *Qu* marks the DP in the subject position, and is not a case marker (Erlewine, to appear).

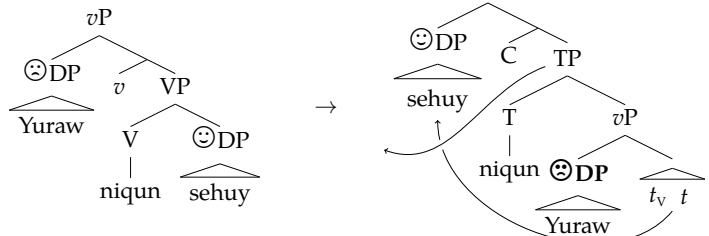
### Non-Actor Voice derivation

In NAV clauses, the actor is **genitive marked**.

- This genitive-marked actor has been analyzed as an **ergative** argument in some previous work (Huang, 1994; Starosta, 1999; Aldridge, 2004).

Here is a Patient Voice example:

- (19) Nìq-un na Yuraw (qu) sehuy.  
 eat-PV GEN Yuraw QU taro  
 'Yuraw eats taro.'



This is precisely the configuration where the actor is given genitive/ergative case, using a particular choice of *v* (Aldridge, 2004; Woolford, 2006; Legate, 2008).

## 6 Strategy 2: oblique case

In Dinka NAV clauses, actors appear in a dedicated case, "**marked nominative**" (Koenig, 2006; Van Urk & Richards, to appear), which is tonally marked:

- (20) a. **Ayén** a-cé cuín cà. *Actor Voice (AV)*  
 Ayen.ABS 3S-PREF.AV food.ABS eat.NF  
 'Ayen has eaten food.'  
 b. Cuín a-cíi **Áyèn** cà. *Patient Voice (PV)*  
 food.ABS 3S-PREF.PV Ayen.NOM eat.NF  
 'Food, Ayen has eaten.'

### "Marked nominative" ≠ ergative

"Marked nominative" is not linked to transitivity or semantic properties of the verb and can be found with unergatives (21) and unaccusatives (22):

- (21) a. **Adít** a-nín. *Adit.ABS 3S-sleep.AV*  
 'Adit is sleeping.'  
 b. Nín **Ádít?** *sleep.AV Adit.NOM*  
 'Is Adit sleeping?'
- (22) a. **Galám** a-cé dhuòŋ. *pen.ABS 3S-PREF.AV break.NF*  
 'The pen broke.'  
 b. Cé **gálám** dhuòŋ? *PREF.AV pen.NOM break.NF*  
 'Did the pen break?'

### "Marked nominative" ≠ nominative or default

"Marked nominative" is the marked case. The absolutive appears in all default contexts, as in citation form and on nominal predicates (Andersen, 1991, 2002).

- (23) Adít e-**dupiòc**.  
 Adit.ABS COP-**teacher.ABS**  
 'Adit is a teacher.'

In addition, "marked nominative" is also assigned by some prepositions (Andersen, 2002):

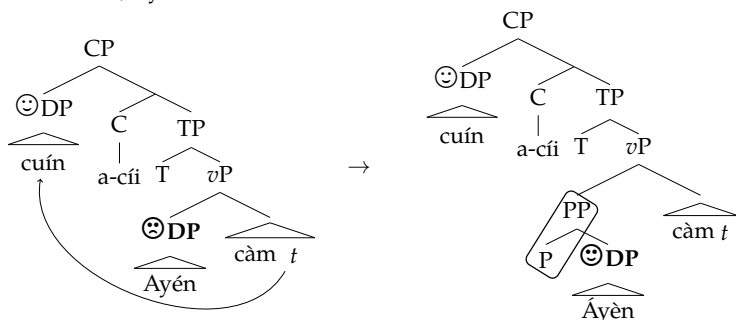
- (24) a. Yín n̄hiàr yòŋ è **Máyèn**.  
 you love.AV house.LNK P **Mayen.NOM**  
 'You love Mayen's house.'  
 b. Cuín a-cíi cà. ne p̄àl ne **Áyèn**.  
 food.ABS 3S-PREF.PASS eat.NF P knife.ABS P **Ayen.NOM**  
 'The food has been eaten with a knife by Ayen.'

## “Marked nominative” as an oblique case repair

**Proposal:** “Marked nominative” is assigned by a null preposition, inserted to license a caseless nominal.

Non-Actor Voice derivation

- (25) Cuín a-cí Àyèn cà.àm.  
 food.ABS 3S-PRF.PV **Ayèn.NOM** eat.NF  
 ‘Food, Ayen has eaten with a knife.’



☞ In this view, “marked nominative” is actually a *prepositional case*, and absolutive is the only real structural case (which is why it behaves like the unmarked case).

For similar proposals regarding prepositions for nominals which would otherwise be unlicensed, see Stowell (1981) on English *of*-Insertion and Halpert (2012) on Bantu augment nominals.

( Strategies 1 and 2 could be seen as very similar, if we analyzed ergative/genitive case in Formosan and Philippine languages (Atayal above) as a last resort repair. See Imanishi (in preparation) for such a proposal for ergativity in Mayan. )

## 7 Strategy 3: pseudo-noun incorporation

A different strategy is found in Balinese. In Balinese, the actor in non-actor voices forms a single “phonological word” with the verb (Clynes, 1995).

**Proposal:** In Balinese NAV clauses, post-verbal actors undergo *Pseudo-Noun Incorporation* (PNI), by means of Morphological Merger (Levin, 2014).

Such actors display **strict head-head adjacency**:

1. Pre-nominal adjectives are banned. Adjectives that can appear pre- and post-nominally appear post-nominally when modifying a non-initial actor (26-a-b)

- (26) a. [(Liu) cicing (liu)] ηgugut Nyoman.  
 (many) dog (many) av.bite Nyoman  
 ‘Many dogs bit Nyoman.’  
 b. Nyoman gugut [(\*)liu) cicing (liu)].  
 Nyoman pv.bite (\*many) dog (many)  
 ‘Many dogs bit Nyoman.’

2. In addition to this, the post-verbal actor shows a definiteness effect. The definite suffix *-e* and overt determiners like *ento* ‘that’ are illicit:

- (27) a. I Wayan gugut cicing.  
 ART Wayan pv.bite dog  
 ‘A dog bit Wayan.’  
 b. \*I Wayan gugut cicing-**e** (**ento**).  
 ART Wayan pv.bite dog-DEF (**that**)  
 ‘The dog bit Wayan.’ (Wechsler & Arka, 1998, p. 441)

We posit that this is because the NPs block PNI of D.

3. In support of this, we see that pronouns and proper names can undergo PNI.

- (28) a. Be-e daar **ida**.  
 fish-DEF pv.eat 3sg  
 ‘(S)he ate the fish.’  
 b. Be-e daar **Nyoman**.  
 fish-DEF pv.eat **Nyoman**  
 ‘Nyoman ate the fish.’

We can derive this if pronouns and proper names occupy D<sup>0</sup> (e.g. Postal, 1966; Longobardi, 1994; Elbourne, 2001) and lack an NP, satisfying head-head adjacency.

## 8 Conclusion

Today we presented two examples where the one-to-one correspondence of voice, case, and extraction can **break down** in voice system languages.

We adopt the idea that voice morphology is a form of **extraction marking**, which tracks the argument moved to the “subject” position (Chung, 1994; Richards, 2000; Pearson, 2001, 2005). By connecting this  $\bar{A}$ -position to the licensing of the actor in AV clauses, we arrive at a unified explanation for **the quirky behavior of actors in NAV clauses**:

1. Ergative/genitive marking in Formosan and Philippine languages;
2. Oblique case marking in Dinka (Nilotic);
3. Pseudo-noun incorporation in Balinese.

Specifically, the idea is that (a) the external argument actor lacks structural licensing in its Spec, $\nu$ P position, (b) the actor is licensed in the subject position in AV, and (c) another strategy is necessary for licensing the subject in NAV clauses.

### Some questions for discussion:

1. What is extraction marking? Cf. Chung (1994), Pesetsky & Torrego (2001), Rizzi & Shlonsky (2007), Erlewine (2014)
2. Why and how do voice languages show more articulated extraction marking, like benefactive voice or oblique voice?

Possible answer: Such voices reflect argument structure alternations (Rackowski, 2002) or preposition-stranding/incorporation (Guilfoyle et al., 1992; Van Urk, in preparation), necessary to turn PPs into nominals that can occupy the mixed “subject” position.

## References

- Aldridge, Edith Catherine. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.
- Aldridge, Edith Catherine. 2008. Phase-based account of extraction in Indonesian. *Lingua* 111:1440–1469.
- Andersen, Torben. 1991. Subject and topic in Dinka. *Studies in Language* 15:265–294.
- Andersen, Torben. 2002. Case inflection and nominal head marking in Dinka. *Journal of African Languages and Linguistics* 23:1–30.
- Chung, Sandra. 1976. An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7:41–87.
- Chung, Sandra. 1994. Wh-agreement and “referentiality” in Chamorro. *Linguistic Inquiry* 25:1–44.
- Clynes, Adrian. 1995. Topics in the phonology and morphosyntax of Balinese - based on the dialect of Singaraja, North Bali. Doctoral Dissertation, Australia National University.

- Cole, Peter, and Gabriella Hermon. 2005. Subject and non-subject relativization in Indonesian. *Journal of East Asian Linguistics* 14:59–88.
- Dimmendaal, Gerrit. 1985. Prominence hierarchies and Turkana syntax. In *Proceedings of ACAL 14*, 127–148.
- Elbourne, Paul. 2001. E-type anaphora as NP-deletion. *Natural Language Semantics* 9:241–288.
- Erlewine, Michael Yoshitaka. 2014. Anti-locality and optimality in Kaqchikel Agent Focus. URL <http://ling.auf.net/lingbuzz/001841>, manuscript, MIT.
- Erlewine, Michael Yoshitaka. to appear. Subject marking on non-subjects in Squliq Atayal. In *Proceedings of AFLA 20*.
- 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.
- Halpert, Claire. 2012. Argument licensing and agreement in Zulu. Doctoral Dissertation, Massachusetts Institute of Technology.
- Huang, Lillian Meijin. 1994. Ergativity in Atayal. *Oceanic Linguistics* 33:129–143.
- Imanishi, Yusuke. in preparation. Doctoral Dissertation, Massachusetts Institute of Technology.
- Koenig, Christa. 2006. *Case in Africa*. Oxford University Press.
- Legate, Julie Anne. 2008. Morphological and abstract case. *Linguistic Inquiry* 39.
- Legate, Julie Anne. 2012. Subjects in Acehnese and the nature of the passive. *Language* 88:495–525.
- Levin, Theodore. 2014. Pseudo noun incorporation is M-Merger: evidence from Balinese. Handout from LSA 2014.
- Longobardi, Giuseppe. 1994. Reference and proper names: a theory of N movement in syntax and logical form. *Linguistic Inquiry* 25:609–665.
- Pearson, Matthew. 2001. The clause structure of Malagasy: A minimalist approach. Doctoral Dissertation, University of California at Los Angeles.
- Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language & Linguistic Theory* 23:381–457.
- Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*. MIT Press.
- Postal, Paul M. 1966. On so-called “pronouns” in English. In *19th monograph on languages and linguistics*. Georgetown University Press.
- Rackowski, Andrea. 2002. The structure of Tagalog: Specificity, voice, and the distribution of arguments. Doctoral Dissertation, Massachusetts Institute of Technology.
- Richards, Norvin Waldemar, III. 2000. Another look at Tagalog subjects. In *Formal issues in Austronesian linguistics*. Kluwer.
- Rizzi, Luigi, and Ur Shlonsky. 2007. Strategies of subject extraction. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. Uli Sauerland and Hans-Martin Gärtner, volume 89 of *Studies in Generative Grammar*. Mouton de Gruyter.
- Rohde, Hannah. 2006. The syntax of questions in Moro. Handout from ACAL 37.
- Rose, Sharon, Farrell Ackerman, George Gibbard, Peter Jenks, Laura Kertz, and Hannah Rohde. 2014. In-situ and ex-situ wh-question constructions in Moro. *Journal of African Languages and Linguistics* 35:91–125.
- Starosta, Stanley. 1999. Transitivity, ergativity, and the best analysis of Atayal case marking. In *Selected papers from the Eighth International Conference on Austronesian Linguistics*.
- Stowell, Timothy Agnus. 1981. Origins of phrase structure. Doctoral Dissertation, Massachusetts Institute of Technology.
- van Urk, Coppe. in preparation. Agreement and the left periphery in Dinka. Doctoral Dissertation, Massachusetts Institute of Technology.
- van Urk, Coppe, and Norvin Waldemar Richards, III. to appear. Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry*
- Wechsler, Stephen, and Wayan Arka. 1998. Syntactic ergativity in Balinese: An argument structure based theory. *Natural Language & Linguistic Theory* 16:387–441.
- Woolford, Ellen. 2006. Lexical case, inherent case, and argument structure. *Linguistic Inquiry* 37.
- Yanti. 2010. A reference grammar of Jambi Malay. Doctoral Dissertation, University of Delaware.