

Two voices calling out as one: A split voice analysis of Javanese passives

Jian Cui & Jack Isaac Rabinovitch*

Abstract. This paper examines the voice system of the Surakarta dialect of Javanese (SkJ) and proposes that both syntactic ergativity and a bundled C/T probe are necessary in interpreting the empirical facts in the language. We present an analysis along these facts, and further propose a decomposed voice structure based on the morphosyntactic realizations of agents in SkJ voice system, which directly feeds the split ergativity analysis that underpins the extraction restriction observed in SkJ.

Keywords. Javanese; Austronesian languages; voice; extraction restriction; split ergativity; A/A' distinction; passives

1. Introduction. A longstanding concern with the Austronesian voice system has been its systematic alternations in word order, case marking, and voice morphology, and how these phenomena interact with A' extraction. Notably, A' extraction in many Austronesian languages is restricted to only subjects or pivots. Many competing analyses have been proposed to account for such restrictions. In A-oriented approaches, voice morphology drives the course of derivation, encodes argument structure alternations which underpin extraction restrictions, and results in case morphology in languages that overtly mark case (e.g., Aldridge 2004, 2008; Legate 2012). A'-oriented approaches focus on the striking similarities between pivots and topics in terms of definiteness restrictions, binding properties, and multiple extractions, and posit that voice morphology is simply a byproduct that marks the result of the extraction (Rackowski 2002; Pearson 2005; Rackowski & Richards 2005; Erlewine et al. 2017; a.o.). Finally, for some languages, which exhibit mixed A/A' properties in their voice systems, linguists argue for a mixed A/A' approach, in which composite probes, formed of bundled C and T heads drives both the A and A' properties of movement to subject/pivot position (Aldridge 2017; Erlewine 2018; Colley & Privoznov 2020).

In this paper, we provide an analysis of the SkJ voice system, which takes aspects of both the traditional A-oriented approach, and mixed A/A' approaches. While we maintain the split-ergativity analysis of traditional A-oriented approaches, we incorporate the concept of a composite probe to better account for the empirical facts observed in SkJ. Movement to pivot position (MPP) in SkJ exhibits distinct mixed A/A' properties, suggesting that the pivot undergoes mixed A/A' movement to an A/A' position, necessitating the integration of the composite C/T head.

In the remainder of Section 1, we describe the core properties of Austronesian voice systems and give a brief background on Javanese. In Section 2, we present novel data from SkJ, illustrating its voice system with a focus on how agents are morpho-syntactically realized in non-actor voices (NAV). We argue the facts are best analysed as having a decomposed verb phrase structure. Section 3 examines the profile of extraction restriction in SkJ. We provide comprehensive evidence showing that the movement to pivot position in the language exhibits mixed A/A' properties, suggesting the involvement of a composite probe that bundles C and T heads. We then argue that the extraction profile is best accounted for under syntactic ergativity. In Section 4, we discuss the implications of this study.

* We would like to acknowledge our consultant Ismartilah Drummond for her kindness and help with our Javanese data, the WOLF Lab @ Harvard Linguistics for funding and encouraging this project, Tanya Bondarenko and Dmitry Privoznov for their guidance, and the audience at AFLA 30 and LSA 2024 for their comments and suggestions. Authors: Jian Cui, Harvard University (jiancui@fas.harvard.edu) & Jack Isaac Rabinovitch, Harvard University ([jrabinovitch@g.harvard.edu](mailto:jrabovitch@g.harvard.edu)).

1.1. JAVANESE VOICE SYSTEM AND COMPETING ANALYSES. Austronesian-type voice systems stand out in demonstrating four essential properties: i. voice alternations always come along with ‘pivotization’ of a privileged argument — this argument, irrespective of its grammatical function, is usually designated a consistent prominent structural position or realized with a particular morphological marking; ii. voice morphology on the verb depends on which argument is chosen as pivot, forming a system where voice morphology cross-references the pivot; iii. pivot arguments are the only arguments available for subsequent A'-extractions (i.e., topicalization, wh-movement, and relativization); and iv. non-pivot arguments are sometimes morphologically marked as well, and they usually come with a genitive case. Not all languages in the family exhibit this whole constellation of properties. Surakarta Javanese, as a typical Austronesian language, demonstrates a partial profile:

- The **pivot** normally occurs in a preverbal position — ie. the surface subject. It must be interpreted as definite¹.
- The pivot is **strictly cross-referenced** by the voice morphology: Agent pivotization necessitates an actor voice construction; Theme pivotization requires non-actor voice morphology on the verb.
- **A' extractions** exclusively target pivot arguments without any exception.
- **Non-pivot arguments** are not consistently morphologically marked: SkJ lacks overt case.

2. The voice system in Surakarta Javanese (SkJ). Javanese has been traditionally described as possessing a three-way voice alternation: actor voice, object voice and passive voice (Robson 2014; Adelaar & Himmelmann 2005; Connors 2008; Nurhayani 2014; a.o.). In their more recent study on the Surabaya dialect of Javanese (SbJ), Patrianto & Chen (2023) characterize the actor voice as being marked with a homorganic nasal prefix /N-/. In SbJ, object voice is restricted to first or second person singular agents involving proclitics which represent the agent's person and number. Finally, passive voice is distinguished by the verbal prefix *di-* and is limited to third person external arguments, either in bare DP or *by*-phrase forms. Surabaya and Surakarta Javanese differ in a few ways; in this section, we present the picture of Surakarta Javanese, and show that not all Javanese dialects fall within the traditional characterization of Javanese voice as previously discussed, thus requiring a more careful examination. These differences require a reorganization of the Javanese voice system in SkJ, based on more nuanced morpho-syntactic realizations of the agent and the distinct syntactic properties observed across various voice constructions in the language.

Word order. Surakarta Javanese has a basic word order of SOV. In actor voice, this surfaces as *agent-verb-theme* string as the basic order (1).

- (1) Surti nulis suraté.
 Surti ACT.write letter-DEF
 ‘Surti wrote the letter.’

Non-actor voices involve pivot promotion of the theme. As a result, the theme appears in subject/pivot position. The agent may either be omitted, or appear as a PP headed by *dening* ‘by’

¹ Pivots may also be postposed to the right of the rest of the main clause.

(2a) or a DP (2b), in either case appearing generally after the verb. This results in an order of *theme-verb-(agent)*.

- (2) a. Suraté di-tulis (dening Surti).
 letter-DEF NAV-write by Surti
 ‘The letter was written (by Surti).’
 b. Suraté di-tulis Surti.
 letter-DEF NAV-write Surti
 ‘The letter was written by Surti.’

In addition to these basic word orders, there are other word orders which are possible, but require particular syntactic configurations. We will not focus on these other word orders in this paper.

2.1. NON-ACTOR VOICES IN SURAKARTA JAVANESE.

The *tak-/kok-* construction. Like Sbj, SkJ has a construction which consists of proclitics, marking first and second person singular agents, marked *tak-* (1SG) and *kok-* (2SG) respectively (3a). *Tak-* and *kok-* are incompatible with additional agent-denoting bare DPs or *by*-phrases (3b).

- (3) a. Suraté {tak /kok }-tulis.
 letter-DEF {1SG.CL /2SG.CL }-write
 ‘The letter was written by {me / you.SG}.’
 b. *Suraté {tak /kok }-tulis (dening) {aku /kowe}.
 letter-DEF {1SG.CL /2SG.CL }-write by {1SG /2SG}
 ‘The letter was written by {me / you.SG}.’

The *di-* construction with DP agents. agents may be explicitly expressed via postverbal bare DPs, only if agent is neither 1SG nor 2SG (4). The unavailability of 1/2SG pronouns as agents in bare DP constructions cannot be due to their pronominal nature, because we can see that the bare DP can otherwise be a 3SG pronoun in *dhéwékè*.

- (4) Surat-é di-tulis {Surti/ aku saklorong / *aku / *kowe / dhéwékè}.
 letter-DEF NAV-write {Surti/ 1 two / 1SG / 2SG / 3SG}
 ‘The letter was written by {Surti / us two / *me / *you_{SG} / him or her}.’

As a result, we obtain a complementary distribution between *tak-/kok-*, which is compatible to all but 1SG/2SG agents, and bare DP constructions, which is strictly confined to 1SG/2SG agents.

The *di-* construction with PP agents. In *di*-constructions, agents may also be realized overtly as *by*-phrase headed by *dening*, in which case any person/number features are possible (5).²

² Our consultant indicated that the *di-* passive with a *by*-phrase is compatible with 1SG/2SG agents, but is dispreferred outside of narrative or written contexts, with preference given to the *tak-/kok-* construction.

- (5) Surat-é di-tulis dening {Surti / aku / kowe / aku saklorong }.
 letter-DEF NAV-write by {Surti / 1SG / 2SG / 1 two }
 ‘The letter was written by {Surti / me / you_{SG} / us two}.’

When there is no bare DP or *by*-phrase present in the structure, *di*-passive construction contains an implicit agent, which can bind into the PRO in adjuncts.³ In (6), the person who wrote the letter is co-referenced with the person who invite the guest, suggesting the presence of an implicit agent in the clause.

- (6) Surat-é di-tulis kanggo PRO ngundang tamu.
 letter-DEF NAV-write in.order.to PRO ACT.invite guest
 ‘The letter was written (by *x*) in order (for *x*) to invite a guest.’

Optional preposition dropping in *di*-constructions? So far, we have examined all possible non-actor forms in SkJ. One immediate speculation may come to mind: are *di*- constructions with bare DPs simply reduced forms of *di*- constructions with PP agents via preposition dropping? We show four pieces of evidence to exclude this possibility and suggest that a syntactic distinction should be made between the two: i. preposition dropping is generally prohibited; ii. flexible word order is available for *by*-phrases, but not bare DP agents of *di*-constructions; iii. adverbial insertion is only available for *by*-phrase constructions, not bare DP structures; and iv. *by*-phrase agents can undergo wh-extraction whereas bare DP agents of *di*-constructions cannot.

First, preposition dropping is generally disallowed in SkJ. In a ditransitive construction like (7a), the preposition *nèng* ‘to’ cannot be omitted. The only way to get rid of the preposition, just like English, would inevitably involve a change in word order (7b).

- (7) a. Obligatory prepositions
 Surti ng-irim-i surat *(nèng) Tono.
 Surti ACT-send-I letter to Tono
 ‘Surti sent a letter to Tono.’
 b. Word order changes when the preposition is absent
 Surti ng-irim-i Tono surat.
 Surti ACT-send-I Tono letter
 ‘Surti sent Tono a letter.’

Second, certain kinds of alternative word order, such as VSO, are available for *by*-phrase constructions (8a) but not for a bare DP constructions (8b).

- (8) a. di-tulis surat-é dening Esti.
 NAV-write letter-DEF by Esti
 ‘The letter was written by Esti.’
 b. *di-tulis surat-é Esti.
 NAV-write letter-DEF Esti
 Intended: ‘The letter was written by Esti.’

³ Surprisingly, this implicit agent can be further bound by higher quantifiers, being treated like a pronoun. For discussion on this property of implicit agents, and for a semantic account, see Cui & Rabinovitch (to appear).

Third, adverbial insertion is only licensed with a *by*-phrase non-actor form (9a), but not with bare DP or proclitic construction (9b–9c) based on diagnostics in Nurhayani (2014).

- (9) a. Surat-é di-tulis (cêpêt-cêpêt) dening Surti.
 letter-DEF NAV-write quickly by Surti
 ‘The letter was (quickly) written by Surti.’
 b. Surat-é di-tulis (*cêpêt-cêpêt) Surti.
 letter-DEF NAV-write quickly Surti
 ‘The letter was (*quickly) written by Surti.’
 c. Surat-é {tak / kok }-(*cêpêt-cêpêt)-tulis.
 letter-DEF {1SG.CL / 2SG.CL }-quickly-write
 ‘The letter was (*quickly) written by {me / you.SG}.’

Finally, we replicate diagnostics in Nurhayani (2014) that *Wh*-extraction is only allowed for *by*-phrase constructions but not bare DP constructions. In (10a), the *wh*-element *sapa* ‘who’ appears within the *by*-phrase agent of the clause. The *by* phrase *dening sapa* can undergo *wh*-movement into the left periphery (10b)

- (10) a. Suraté di-tulis dening sapa?
 letter-DEF NAV-write by who
 ‘Who wrote the letter?’
 b. Dening sapa suraté di-tulis?
 by who letter-DEF NAV-write
 ‘Who wrote the letter?’

However, when *sapa* is instead generated within a DP agent in a non-pivot position, it must remain in situ, as in (11a), and is unable to extract (11b).

- (11) a. Suraté di-tulis sapa?
 letter-DEF NAV-write who
 ‘Who wrote the letter?’
 b. *Sapa sing di-tulis suraté?
 who REL NAV-write letter-DEF
 ‘Who wrote the letter?’

Based on these four diagnostics, we argue that the bare DP and *by*-phrase *di*- constructions should be teased apart in syntax, rather than seeing the former as a reduced form of the latter.

Upon analyzing SkJ non-actor voice constructions along with their respective syntactic characteristics and agent profiles, we arrive at these primary generalizations:

- A **complementary distribution** is yielded between the bare DP *di*- configuration, which is compatible to all but 1SG/2SG agents, and the *tak*-/ *kok*- construction, which is strictly confined to 1SG/2SG agents. This implies an intrinsic identity of the two forms and calls for a unification in syntax despite their distinct voice morphology (Configuration 1).

- A **syntactic distinction** is implied between the bare DP and the *by*-phrase *di*- constructions (Placing *by*-phrases into a distinct Configuration 2).
- A **syntax-morphology mismatch** is yielded in (1), where the split of voice morphology does not correspond perfectly with the syntactic division that we would otherwise propose.

Agent Realization	Agent Restriction	Morphology	Syntax
proclitic <i>tak-/kok-</i>	1sg/2sg	<i>tak-/kok-V</i>	Configuration 1
Bare DP	all but 1sg/2sg	<i>di-V</i>	
<i>by</i> -phrase (PP)	all person/numbers		Configuration 2
Implicit Agent	—		

Table 1. Summary on SKJ non-actor voices

2.2. ANALYSIS. In this section, we argue for a reanalysis of SkJ voice system that can make sense of the two predicted non-actor voice syntactic configurations, as well as their apparent mismatch with morphological exponence. In Section 2.2.1, we argue that SkJ has two non-actor voices: a simple passive, which consists of constructions with optional *by*-phrase agents (Configuration 2 in example 1), and a split object voice that is built on top of the syntactic structure of the simple passive, and encompasses both *tak-/kok-* and bare DP constructions (Configuration 1 in example 1). In Section 2.2.2, we argue that the mismatch between syntax and morphology arises from a combination of spanning exponence rules in SkJ, and the fact that the object voice syntactically ‘contains’ the simple passive.

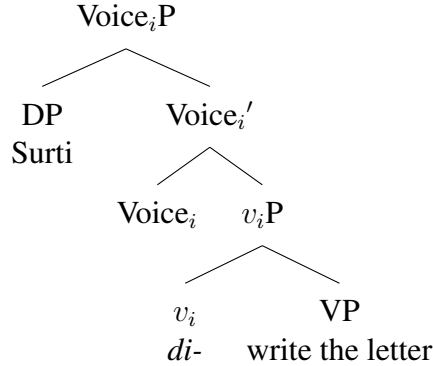
2.2.1. A SPLIT-VOICE ANALYSIS. We follow Harley (2013); Wurmbrand (2021) in assuming that the voice domain may be composed into multiple parts, specifically, we propose that the SkJ voice system comprises of a three-layered verb phrase structure.

- **Lexical VP**: introduces selected internal argument
- **Verbalizing v_i P**: only semantically refers to a pronominal agent
- **Functional projection VoiceP**: overt agent arguments are syntactically introduced

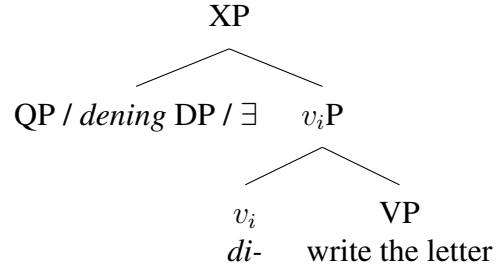
The two different non-actor voice configurations arises from the optional projection of Voice — passive voice (implicit and *by*-phrase agents) contains the Lexical VP and Verbalizing v_i P but lacks VoiceP; object voice (bare DP and *tak/kok-* agents) is built on top of the passive voice, including a Voice $_i$ P layer.⁴ Ignoring the *tak/kok-* constructions for now, we can illustrate object voice as follows: in (12a), the v_i attaches to the VP, and surfaces as the passive morpheme *di-*. The resulting v_i P merges with a Voice $_i$ head, which functions to provide an open argument position for the agent (Cui & Rabinovitch to appear). This position must then be filled by a DP, such as the name *Surti*, which is then interpreted as the agent. When the agent is implicit or realized as a PP *by*-phrase, no VoiceP is projected, and so there is no position for a bare DP to fill. Rather, the agent must remain implicit, or can surface in the form of a *by*-phrase (12b).

⁴ Following Cui & Rabinovitch (to appear), We adopt the use of indices on non-actor v and Voice heads to model the pronominal nature of agents in SkJ non-actor voices.

(12) a. **Bare DP (not 1SG/2SG)**



b. **Implicit Agent/By-Phrase**



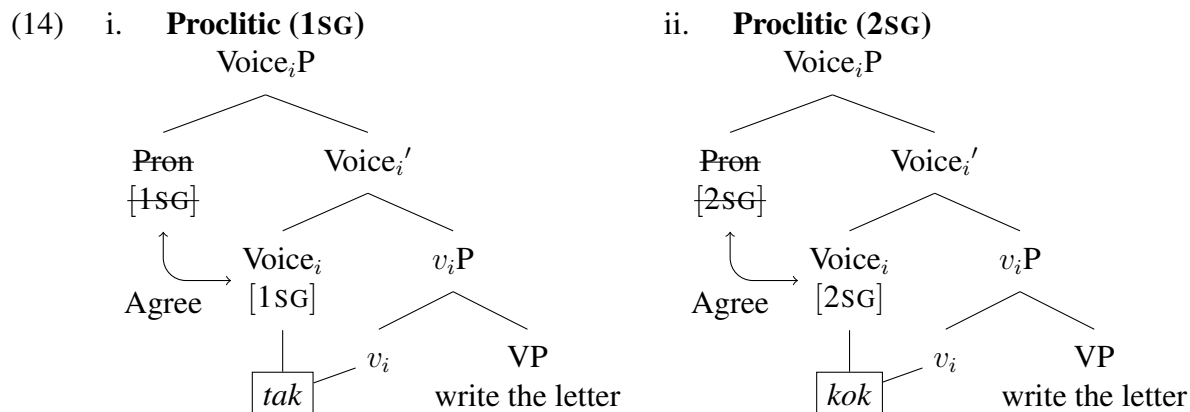
2.2.2. RESOLVING THE SYNTAX-MORPHOLOGY MISMATCH. Object voice, as proposed above, subsumes not only bare DP constructions but also *tak-/kok-* structures, featuring a complementary distribution in their agent profiles. First, as illustrated in (14), the same decomposed verb phrase structure is posited for *tak-/kok-* configurations, which are deemed as underlying instances of bare DP constructions with 1/2SG pronouns occupying the Spec, VoiceP. Following this structure, the derivation of the surface complementary distribution involves three principal components: chain reduction, span formation, and allomorph competition.

Chain reduction. We assume that Voice_i involves a probe which searches for local singular participant $[+PART, +SG]$ DPs. When first and person singular pronouns merge, they undergo agreement and share all their features with Voice_i . Following Holmberg et al. (2009) and Rabinovitch (2022), because the agent pronoun forms a subset of the features of Voice_i , the two heads undergo chain reduction, where the pronoun deletes at PF. As a result, Voice_i contains first/second person features and singular features (14).

Span formation. We assume the following span-based spell-out rules: the indexed v_i , shared by both object and passive voices, has a default spell-out as */di-/*. However, when Voice_i carries with it the appropriate set of features, it can form a span with v_i , surfacing as *tak-* (when Voice_i carries 1SG features) or *kok-* (when Voice_i carries 2SG features).

(13) Three spell-out rules

- a. $\langle v_i \rangle \rightarrow /di-/$
- b. $\langle v_i, \text{Voice}_i, 1SG (= [+PART, +AUTH, +SG]) \rangle \rightarrow /tak-/$
- c. $\langle v_i, \text{Voice}_i, 2SG (= [+PART, -AUTH, +SG]) \rangle \rightarrow /kok-/$



Allomorph competition. In object voice, the formation of spans creates allomorphic variation in voice morphology, which is conditioned by the features of the agent in the Spec, VoiceP position. When sufficient features required for *tak-/kok-* constructions are present (v_i and Voice_{*i*} with 1SG/2SG features), competition between the *di-* and *tak-/kok-* spell-outs favor the more specific *tak-/kok-*. Thus, when first or second person pronouns merge at SpecVoice_{*i*}P, they are systematically deleted, and their features are shared with Voice_{*i*}, triggering span formation, resulting in the complementary distribution between *tak-/kok-* and *di-* constructions with bare DP agents.

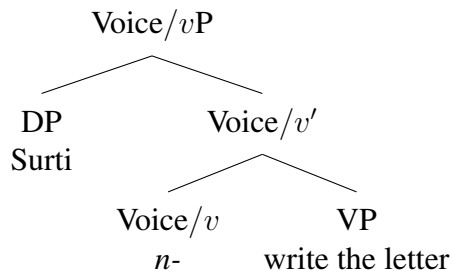
Summary. Non-actor voices in Javanese are split between two syntactic configurations: the passive voice, involving a v P projection, and object voice, involving the larger VoiceP projection. Both of these projections include v_i , which surfaces as *di-* in most contexts; due to a combination of chain- and span-forming processes in SkJ, v_i may form part of a span, surfacing as a proclitic *tak-* or *kok-*. These environments require specific features on Voice_{*i*}, which result from agreement with first and second singular pronouns in SpecVoice_{*i*}P. The result is that, while both passive and object voices can surface with *di-* morphology, object voice surfaces with proclitics when the agent is interpreted as first or second person singular, giving rise to an apparent mismatch between syntax and morphology.

2.3. MODELLING ACTOR VOICE. In SkJ, actor voice is marked by the nasal prefix *N-* and requires an overt DP agent, and involves no promotion of the theme.⁵ Unlike object voice, where v and voice are distinct and the external argument is introduced in Spec, VoiceP, the actor voice (active) sentence (15) features the agent, *Surti*, introduced in the specifier position of a phrase projected by a *bundled* Voice/ v head (2.3).⁶ We follow Cui & Rabinovitch (to appear) in assuming this bundled head has similar semantics to a conventional agent introducing agentive v . Thus, in SkJ, both actor and object voices involve Voice and v_i heads. However, they are distinguished by whether these heads are bundled together.

⁵ *N-* tends to assimilate to the place features of the following consonant.

⁶ We thank Yining Nie for suggesting a bundled analysis for actor voice.

- (15) Surti nulis suraté.
 Surti ACT.write letter-DEF
 ‘Surti wrote the letter.’



3. Movement to Pivot Position (MPP) and A' extraction. We now shift our focus to the examination of Movement to Pivot Position (MPP) in SkJ non-actor voice constructions. In SkJ, MPP exhibits both A and A' properties. For this paper, we demonstrate i. that MPP forms new antecedents for binding — something expected for A-movement, and ii. that MPP is required for a DP to undergo relativization, suggesting that the landing site of MPP is an A' position.⁷

3.1. MPP FORMS NEW ANTECEDENTS FOR BINDING. In actor voice, agents asymmetrically c-command themes, preventing themes from binding anaphora in the agent's position (16a).

- (16) a. *Dhèwèké dhèwèk ngamplèng Surti.
 3SG.REFL AV-punch Surti
 Intended: 'Herself_i punched Surti_i.'

In non-actor voices, the theme moves into pivot position. If MPP were purely A' we would expect that the pivot position cannot form a new antecedent, and so the theme should still be unable to bind the agent. However, pivot themes can bind the agent, both when the agent is a bare DP (17a) or a *by*-phrase (17b), suggesting that the theme undergoes A-movement to Pivot position.

- (17) a. Surti di-kamplèng dhèwèké dhèwèk.
 Surti NAV-punch 3SG.REFL
 'Surti_i was punched by herself_i.'
 b. Surti di-kamplèng dening dhèwèké dhèwèk.
 Surti NAV-punch by 3SG.REFL
 'Surti_i was punched by herself_i.'

3.2. PIVOTS ARE TARGETED FOR A'-EXTRACTION. Many Austronesian languages observe the 'extraction restriction,' a phenomena in which A' operations are unable to target arguments other than the pivot for movement. Such is the case with SkJ. Here we discuss relativization, which cannot target non-pivot arguments in SkJ. For instance, agents in SkJ can only be relativized from within actor voice clauses, where they are the pivot (18a), but not from within object voice clauses, where they are not (18b).

⁷ Other properties we do not have space to discuss includes: MPP circumvents weak crossover effects and precludes reconstruction (A properties); the pivot is restricted to being a DP (A property); pivots must be interpreted as definite (A' property), and pivots are targeted for topicalization and wh-movement (A' property).

- (18) a. Surti wêruh asu-né [_{AV} sing — mangan pisa-né].
 Surti see dog-DEF REL AV.eat pizza-DEF
 ‘Surti saw the dog that ate the pizza.’
 b. *Surti wêruh asu-né [_{OV} sing pisa-né di-pangan —].
 Surti see dog-DEF REL pizza-DEF NAV-eat
 Intended: ‘Surti saw the dog that the pizza was eaten by.’

Themes, on the other hand, cannot be relativized from within actor voice clauses, as in (19a). Instead, they can be relativized from object (19b) and passive (19c) voice clauses, where they serve as pivots. The overarching generalization regarding relativization patterns is that A' extraction selectively targets pivots, which is then reflected in the voice morphology of the language.

- (19) a. *Surti wêruh pisa-né [_{AV} sing asu-né mangan —].
 Surti see pizza-DEF REL dog-DEF AV.eat
 Intended: ‘Surti saw the pizza that the dog ate.’
 b. Surti wêruh pisa-né [_{OV} sing — di-pangan asu-né].
 Surti see pizza-DEF REL NAV-eat dog-DEF
 ‘Surti saw the pizza that was eaten by the dog.’
 c. Surti wêruh pisa-né [_{PV} sing —] di-pangan dening asu-né].
 Surti see pizza-DEF REL NAV-eat by dog-DEF
 ‘Surti saw the pizza that was eaten by the dog.’

3.3. AN ACCOUNT OF MIXED A/A' MOVEMENT. We follow the featural view of the A/A'-distinction (Van Urk 2015), where A-movement is motivated by a pure ϕ probe, which attracts the closest element that has a ϕ feature, whereas A'-movement is triggered by a probe that carries pure A' features (e.g., [+TOP], [+WH]). Movement with mixed properties are caused by a composite probe, which seeks a goal that carries both A and A' features.

We propose that in SkJ, C and T heads are bundled together as C/T (see also Martinović 2015, 2017, 2022; Erlewine 2018, 2020; Colley & Privoznov 2020), integrating ϕ features from T and A' from C, functioning as a composite probe [ϕ , A'] which targets DPs with sufficient A' features. C/T inherits the case-licensing property from T and assigns case to its specifier (i.e., a position that pivot moves into), as well as A' features which probe for definite DPs, deriving the A' properties (i.e., definiteness and extraction restrictions) on pivots. Building upon this, we further propose that SkJ exhibits split case alignment, involving two different C/T-heads:

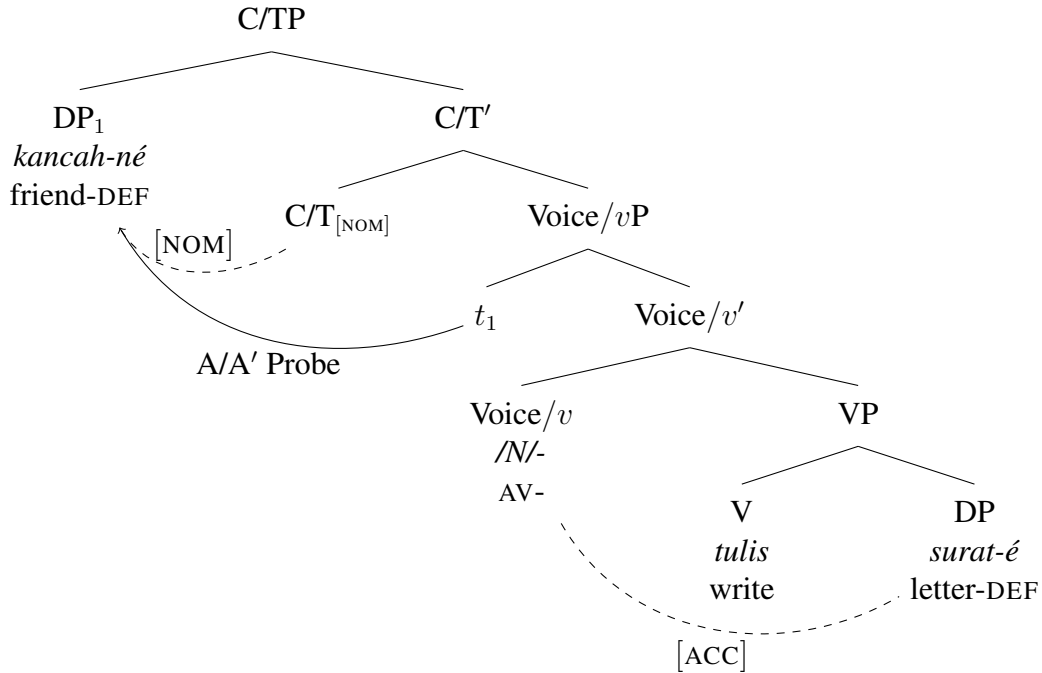
- C/T_[NOM] licenses [NOM] case and selects *v*P or a bundled Voice/*v*P — bundled Voice/*v* licenses [ACC].
 - Prediction: actor and passive voices (*di-* construction with optional *by*-phrase agents) with non-split voice domains involve NOM-ACC alignment.
- C/T_[ABS] licenses [ABS] and selects VoiceP — Voice_{*i*} licenses [ERG].
 - Prediction: Object voice (bare DP *di-* and *tak/kok*-constructions) with a split verb phrase structure involves ERG-ABS alignment.

The split ergativity analysis of the SkJ voice system is a direct consequence of the split voice proposal for the verb phrase structure, as proposed in Section 2. This analysis reveals that the distinction between ergative and nominative alignments arises from the presence or absence of an additional, independent VoiceP projection above the *v*P layer. As a result, the extraction restriction in A' constructions can be straightforwardly interpreted as A' extraction targeting arguments with an unmarked case, namely [ABS] or [NOM].

3.3.1. DERIVING ACTOR VOICE. Actor voice involves $C/T_{[NOM]}$, which licenses nominative case, and Voice/*v*, which licenses accusative case and has a [D*] feature, requiring its specifier be occupied (20). An example derivation is provided in (21).

(20)	Structure:	$[C/TP \quad C/T_{[NOM]} \quad [_{vP} \quad \text{Voice}/v \quad [VP]]]$
	Spellout:	<i>n-</i>
	EPP:	— [D*]
	Case:	[NOM] [ACC]
	Probe:	$[\phi : _ , A']$

- (21) *Kancah-né n-ulis surat-é.* [actor voice]
 friend-DEF AV-write letter-DEF
 ‘The friend wrote the letter.’

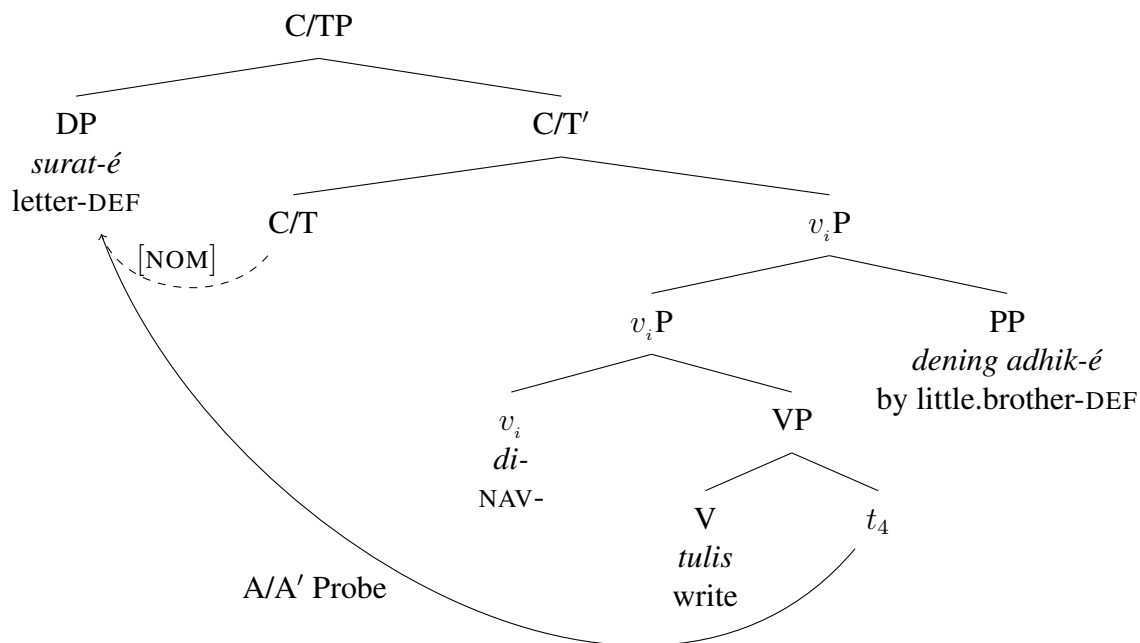


In (21), the agent *kancah-né* fills SpecVoice/*v*P, satisfying Voice/*v*'s [D*] feature. Voice/*v* provides accusative case to the theme *surat-é*. $C/T_{[NOM]}$ probes for the closest DP which satisfies both A and A' properties, agrees with *kancah-né*, moves it into its specifier (pivot position), and gives it nominative case. This pivot argument, bearing an unmarked NOM case and A'-features, can be targeted by subsequent A' extraction to a higher A' position. Additionally, we assume that SkJ involves V-to-T movement, in this case applying by cyclically moving V to Voice/*v* to C/T.

(22)	Structure:	[_{C/TP}	C/T	[_{NOM}	[_{vP}	<i>v_i</i>	[VP]]]]
	Spellout:					<i>di-</i>	
	EPP:		—			—	
	Case:		[_{NOM}				
	Probe:		[<i>ϕ</i> : —, A']				

(22)	Structure:	[_{C/TP}	C/T	[_{NOM}	[_{vP}	<i>v_i</i>	[VP]]]]
	Spellout:					<i>di-</i>	
	EPP:		—			—	
	Case:		[_{NOM}				
	Probe:		[<i>ϕ</i> : --, A']				

(23) surat-é di-tulis dening adhik-é. [passive voice]
 letter-DEF NAV-write by little.brother-DEF
 ‘The letter was written by the little brother.’



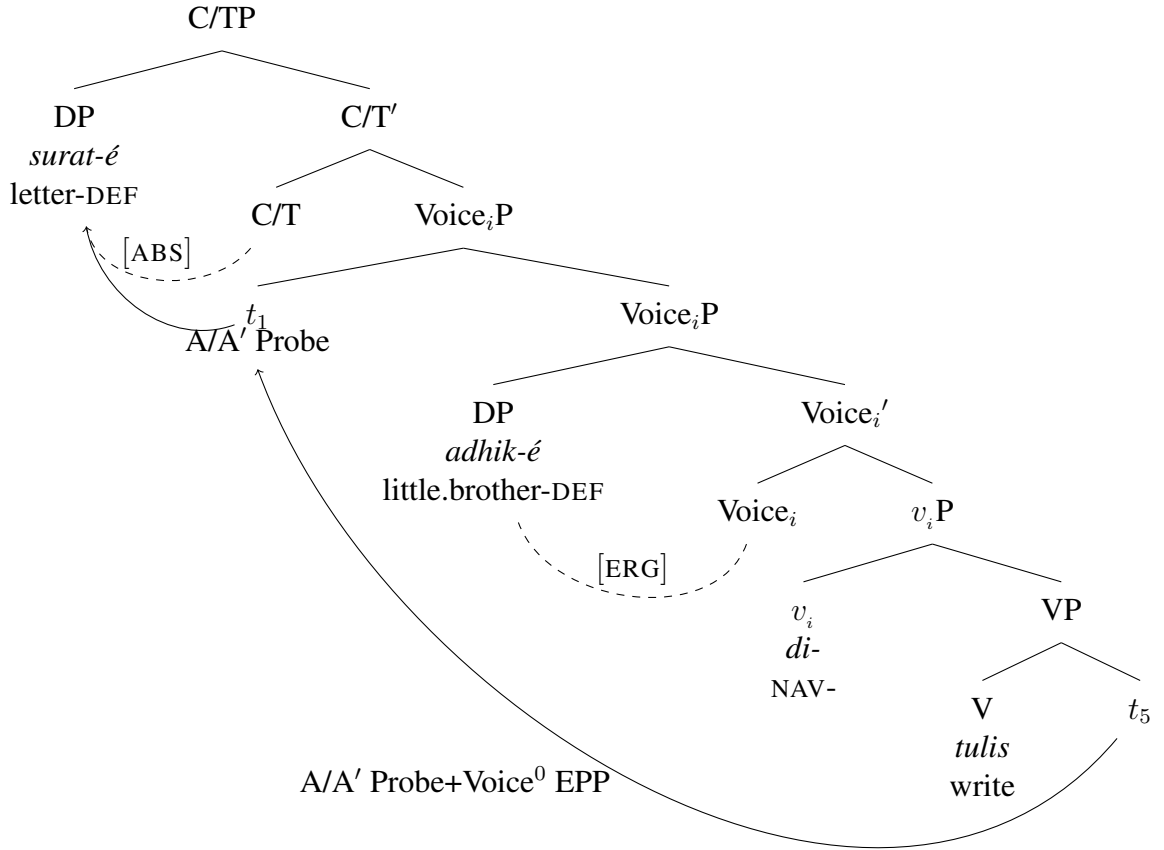
The C/T_[NOM] head scans its c-command domain, identifies the theme — *suraté* — as the closest matching goal and moves it to Spec, C/T_[NOM] where it receives NOM case. Similar to the pattern observed in actor voice, subsequent A' extractions target this unmarked argument. This analysis supports the phase theories of Chomsky (2001) and Aldridge (2004, 2008) — we analyze the bundled Voice/*v* in actor voice as heading a phase whose boundary disallows the C/T head to probe further into the phase for the theme.⁸ An independent indexed *v_i* does not head a phase, and so the theme can be directly pivotized by C/T_[NOM] without passing through Spec*v_i*P.

⁸ This condition in actor voice does not appear to be crucial, as the theme is not a qualifying goal due to the absence of an *obligatory* DEF interpretation.

3.3.3. DERIVING OBJECT VOICE. Object voice involves an independent v_i head, a Voice_i head, and a $\text{C/T}_{[\text{ABS}]}$ head. v_i neither licenses case nor possesses a $[\text{D}^*]$ feature; Voice_i licenses ergative case and hosts a $[\text{D}^*]$ feature and a probe (the same which agrees with pronominal agents); $\text{C/T}_{[\text{ABS}]}$ licenses ablative case, selects the VoiceP , and otherwise is identical to $\text{C/T}_{[\text{NOM}]}$ (24). An example derivation is provided in (25).

(24)	Structure:	$[\text{C/TP}$	$\text{C/T}_{[\text{ABS}]}$	$[\text{VoiceP}$	Voice_i	$[_{vP}$	v_i	$[\text{VP}]]]$
	Spellout:						<i>di/tak/kok-</i>	
	EPP:		—		$[\text{D}^*]$		—	
	Case:		$[\text{ABS}]$		$[\text{ERG}]$			
	Probe:		$[\phi : _ , \text{A}']$		$([\phi : _ , \text{A}'])$			

- (25) surat-é di-tulis adhik-é. [object voice]
 letter-DEF NAV-write little.brother-DEF
 ‘The letter was written by the little brother.’



The $[D^*]$ feature on Voice_i is checked by merging the bare DP agent at its specifier position — the Voice head then assigns the $[\text{ERG}]$ to this argument. Subsequently, the $C/T_{[\text{ABS}]}$, selecting VoiceP, probes into its c-command domain to find the closest matching DP (i.e., the theme *suraté*), extracts this goal to $\text{Spec}, C/T_{[\text{ABS}]}P$ through cyclic movement via the Voice phase edge, and licenses $[\text{ABS}]$. Again, the element that is targeted by A' -extraction is the unmarked (ABS) argument in pivot position ($\text{Spec}C/TP$). Finally, head movement ensures that *di-tulis* appears in C/T , to the left of the agent *adhik-é*, forming the *theme-verb-agent* order of (25).

There are two underlying assumptions behind the process proposed above and illustrated in (25): i. certain probes exhibit case discrimination during the Agreement process (cf., Bobaljik 2008): the agent argument with an ERG case is discriminated when $C/T_{[\text{ABS}]}$ is probing downwards — only the internal argument is probable; ii. Voice with a $[D^*]$ feature, just as Voice/v , serves as a phase and requires cyclic movement of the internal argument to pass through its phase edge. Such cyclic movement is enabled by an (optional) composite probe on Voice, which is only active when Voice is not bundled with v (cf. the Voice/v does not hold such a composite probe).

4. Conclusion. This paper forms a case study on the voice system of Surakarta Javanese. We argue for a need to recategorize actor, passive, and object voice in Surakarta Javanese based on syntactic facts: bare DP and *by*-phrase *di-* constructions should be syntactically distinguished, grouping the former with *tak-/kok-* constructions as split object voice and treating the latter as a simple passive. The resulting mismatch between surface syntax and morphology is explained as morphological exponence deriving a difference between *tak/kok-* and *di-* structures, with *tak/kok-* acting as a span whose derivation induces complementary distribution. We present a proposal for the verb phrase structure in SkJ, building upon previous studies focused on structuring the voice domain (Chomsky 1995; Harley 1995, 2013, 2017; Pylkkänen 2002, 2008; Coon & Preminger 2011). We contend that both voice-bundling and non-voice bundling can coexist within a single language, and their occurrence is conditioned by voice alternations in SkJ: split object voice is built on top of the syntax of passive voice, while actor voice, is uniquely characterized by a bundled form of Voice and v , which is semantically equivalent to a conventional v_{agentive} . With respect to MPP, we propose that the mixed A/A' properties of pivots in SkJ calls for a composite C/T probe. The merged C/T heads select for either VoiceP or vP complements, conditioned by its case features. C/T s which select for VoiceP complements license ABS case, and Voice_i licenses ERG case, deriving the ERG -ABS alignment of split object voice. C/T s which select for vP complements license NOM case, resulting in the NOM-ACC alignment of actor and passive voices.

The phasehood distinction posited in this paper is bolstered by the semantic distinction between v_i and Voice/v and the decomposed verb phrase structure proposed in this analysis. We make the following cross-linguistic predictions: i. eligibility for being a phase is determined by whether the head can introduce an argument in its specifier position — both Voice/vP in actor voice and the VoiceP in object voice form a phase; ii. once a v head can potentially involve an additional, independent projection above it in the voice domain, it loses its eligibility as a phase. Given that non-actor v_i can potentially be selected by a higher Voice in the language, a single v_iP projection in passive voice does not serve as a phase.

What does this study mean for voice and Austronesian in general? In SkJ, we modelled that two distinct voices may call out as one: Object and passive voices may be distinguished syntactically even if they involve the same exponence *di-*. The presence or absence of an independent Voice projection above v can be a source of *both* differences in both voice and case-alignment.

Finally, the Surakarta dialect of Javanese with an A/A'-oriented voice system, should be distinguished from the Surabaya Javanese, featuring a pure A'-oriented voice system, in terms of both syntax and morphology. This suggests that 'a cline of voice systems in transition from a topic-oriented to a subject-oriented system' (Patrianto & Chen 2023) not only occurs within the Austronesian family at large but also within the diverse dialects of Javanese.

References

- Adelaar, K. Alexander & Nikolaus Himmelmann (eds.). 2005. *The Austronesian languages of Asia and Madagascar* (Routledge Language Family Series). London/New York: Routledge.
- Aldridge, Edith. 2004. *Ergativity and word order in Austronesian languages*. Ithaca, NY: Cornell University dissertation.
- Aldridge, Edith. 2008. Phase-based account of extraction in Indonesian. *Lingua* 118(10). 1440–1469. <https://doi.org/10.1016/j.lingua.2007.08.006>.
- Aldridge, Edith. 2017. ϕ -feature competition: A unified approach to the Austronesian extraction restriction. *Chicago Linguistic Society (CLS)* 52. 1–20.
- Bobaljik, Jonathan David. 2008. Where's phi? Agreement as a post-syntactic operation. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi-Theory: Phi features across interfaces and modules*, 295–328. Oxford: Oxford University Press.
- Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Colley, Justin & Dmitry Privoznov. 2020. On the topic of subjects: Composite probes in Khanty. In Mariam Asatryan, Yixiao Song & Ayana Whitmal (eds.), *Proceedings of the North East Linguistics Society (NELS)*, vol. 50, 111–124. UMass Amherst: GLSA Publications.
- Connors, Thomas J. 2008. *Tengger Javanese*. New Haven, CT: Yale University dissertation.
- Coon, Jessica & Omer Preminger. 2011. Transitivity in Chol: A new argument for the Split VP Hypothesis. In *Proceedings of the North East Linguistics Society (NELS)*, vol. 41, UMass Amherst: GLSA Publications.
- Cui, Jian & Jack Rabinovitch. To appear. Indices in the Voice domain: A unified analysis of Javanese passives. *Austronesian Formal Linguistics Association (ALFA)* 13.
- Erlewine, Michael Yoshitaka. 2018. Extraction and licensing in Toba Batak. *Language* 94(3). 662–697. <https://doi.org/10.1353/lan.2018.0039>.
- Erlewine, Michael Yoshitaka. 2020. Anti-locality and subject extraction. *Glossa: A Journal of General Linguistics* 5(1). 84. <https://doi.org/10.5334/gjgl.1079>.
- Erlewine, Michael Yoshitaka, Theodore Levin & Coppe van Urk. 2017. Ergativity and Austronesian-type voice systems. In Jessica Coon, Diane Massam & Lisa Demena Travis (eds.), *The Oxford handbook of ergativity*, 373–396. Oxford: Oxford University Press.
- Harley, Heidi. 1995. *Subjects, events, and licensing*. Cambridge, MA: MIT dissertation.
- Harley, Heidi. 2013. External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua* 125. 34–57. <https://doi.org/10.1016/j.lingua.2012.09.010>.
- Harley, Heidi. 2017. The “bundling” hypothesis and the disparate functions of little v. In Roberta D'Alessandro, Irene Franco & Ángel J. Gallego (eds.), *The verbal domain*, 3–28. Oxford: Oxford University Press.

- Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63(1). 59–97. <https://doi.org/10.1111/j.1467-9582.2008.01154.x>.
- Legate, Julie Anne. 2012. Subjects in Acehnese and the nature of the passive. *Language* 88(3). 495–525. <https://doi.org/10.1353/lan.2012.0069>.
- Martinović, Martina. 2015. *Feature geometry and head-splitting: Evidence from the morphosyntax of the Wolof clausal periphery*. Chicago, IL: The University of Chicago dissertation.
- Martinović, Martina. 2017. Head-Splitting at the Wolof clausal periphery. In Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel & Edward J. Rubin (eds.), *Proceedings of the 34th West Coast Conference on Formal Linguistics*, 364–371. Somerville, MA: Cascadilla Press.
- Martinović, Martina. 2022. Feature geometry and head splitting in the Wolof clausal periphery. *Linguistic Inquiry* 54(1). 79–116. https://doi.org/10.1162/ling_a_00447.
- Nurhayani, Ika. 2014. A unified account of the syntax of valence in Javanese. Ithaca, NY: Cornell University dissertation.
- Patrianto, Hero & Victoria Chen. 2023. Rethinking Indonesian-type passives: Lessons from Javanese. *North East Linguistics Society (NELS)* 53.
- Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language & Linguistic Theory* 23(2). 381–457. <https://doi.org/10.1007/s11049-004-1582-7>.
- Pylkkänen, Mariliina. 2002. *Introducing arguments*. Cambridge, MA: MIT dissertation.
- Pylkkänen, Mariliina. 2008. *Introducing arguments* (Linguistic Inquiry Monograph 49). Cambridge, MA: MIT Press.
- Rabinovitch, Jack Isaac. 2022. Narrow scoping content question items in shifty contexts: A case of surprising non-quotation in Uyghur. *Proceedings of the Linguistic Society of America (PLSA)* 7(1). 5235. <https://doi.org/10.3765/plsa.v7i1.5235>.
- Rackowski, Andrea. 2002. *The structure of Tagalog: Specificity, voice, and the distribution of arguments*. Cambridge, MA: MIT dissertation.
- Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36(4). 565–599. <https://doi.org/10.1162/002438905774464368>.
- Robson, S. O. 2014. *Javanese grammar for students: A graded introduction* (Monash Asia Series). Clayton, AU: Monash University Publishing.
- van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study of Dinka Bor*. Cambridge, MA: MIT dissertation.
- Wurmbrand, Susanne. 2021. A decomposed notion of Voice. Presented at Typology of Cumulativity, Humboldt-Universität zu Berlin, June 2021