THE JAVANESE FREE EXCEPTIVE AND ITS IMPLICATIONS FOR ELLIPSIS CONSTRAINTS IN AUSTRONESIAN*

Luané Lennox
McGill University
luane.lennox@mail.mcgill.ca

Victoria Chen
Victoria University of Wellington
victoria.chen@vuw.ac.nz

Eric Potsdam
University of Florida
potsdam@ufl.edu

There remains considerable debate over the analysis of voice in Western Austronesian languages, whether it is like voice in English and other Indo-European languages or not. Merchant (2013) notes that voice mismatch under clausal ellipsis is impossible in languages with English-type voice. We show using free exceptives and sluicing that this is not the case with Philippine-type voice, where voice mismatches are possible. We take this contrast to argue that Philippine-type voice should not be analyzed in the same way as English-type voice, and is better viewed as morphology hosted outside of the the ellipsis site (i.e. VoiceP), as suggested by recent Ā-approaches to Philippine-type languages. The proposal makes largely correct predictions for Austronesian-languages with European-type voice constructions, where voice mismatch again becomes impossible. We also use this diagnostic to argue that the voice system in Javanese should be analyzed as being of the Philippine-type.

1. Introduction

Relying heavily on Indo-European languages, Merchant (2001, 2013) argues that voice mismatch with a prototypical English-style voice system is impossible under clausal ellipsis. To illustrate, in German, which has an active-passive voice contrast similar to English, (1) and (2) show that voice mismatch under sluicing, a type of clausal ellipsis, results in unacceptability. In (1), the antecedent is in the active voice but the elided clause must be passive given that the case marking on the wh-phrase *wer* 'who.NOM' is nominative, corresponding to the passive subject. Sluicing is impossible. Elided material is enclosed in angled brackets, <...>.[1]

(1) *Voice mismatch in German sluicing – Active-Passive*
*Erika hat jemanden ermordet, aber sie wissen nicht, wer <von ihr
Erika has someone murdered.ACT but they know not who.NOM by her
ermordet wurde>.
murdered.PASS was
(Lit. 'Erika murdered someone, but they don't know who <was murdered by her>.')
(Merchant 2013:81)

*This material is based upon work supported by the National Science Foundation under Grant BCS-2116343, the Marsden Standard Grant 23-VUW-014, and a VUW Research Establishment Grant #410158. Thank you also to the Post-Graduate Student Society of McGill University for travel funding. We thank the following speakers for sharing their language with us: Javanese: Hero Patrianto, Eva Winanta; Acehnese: Muhammad Faris Ridhatillah, Julia Farlia; Tagalog: Ivan Bondoc, Kyla Venice Lerio, Jed Pizarro-Guevara; Malagasy: Voara and Bodo Randrianasolo; Indonesian: Hero Patrianto; Puyuma: Atrung Kagi. Thanks also to Shin Fukuda, Hero Patrianto, Masha Polinsky, Jozina vander Klok, Junko Shimoyama, and the audience at *Malagasy in Montréal* (McGill, 2023) for feedback and comments on earlier versions of the project. We also thank Dan Brodkin for sharing Mandar data with us.

[1] Glossing follows Leipzig glossing conventions with the following additions: AV: actor voice; C: complementizer; PV: patient voice; LV: locative voice; OV: object voice; RED: reduplicant.

In (2), which is also ungrammatical, the antecedent clause is passive but the elided clause must be active because of the nominative case marking on *wer* 'who.NOM'.

(2) *Voice mismatch in German sluicing – Passive-Active*
*Peter wurde von jemandem ermordet, aber sie wissen nicht, wer <ihn
Peter was by someone murdered.ACT but they know not who.NOM him
ermordet hat>.
murdered.PASS has
(Lit. 'Peter was murdered by someone, but they don't know who <murdered him>.')
(Merchant 2013:82)

Free exceptives in many languages, including English, have been argued to be a clausal ellipsis construction, as in (3). It is therefore unsurprising that voice mismatch is also disallowed here. The English examples in (4) show that the voice of the elided clause cannot mismatch with that of the antecedent clause.

(3) Nobody left, except Mary <left>.

(4) a. *Everyone helped me, except by Mary <I was not helped>. [Active-Passive]
b. *I was helped, except Mary <did not help me>. [Passive-Active]

In contrast, it has been occasionally noted that some Philippine-type Austronesian languages allow voice mismatch in clausal ellipsis. Example (6) illustrates this observation for Malagasy sluicing (Potsdam 2007, 2022). The antecedent is in the actor voice but the elided clause is in a non-actor voice.

(5) *Voice mismatch in Malagasy sluicing: Actor voice - Patient voice*
Nandoko zavatra i Bao fa hadinoko hoe inona <no nolokoin' i Bao>.
painted.AV thing Bao but forget.1SG COMP what FOC painted.PV Bao
'Bao painted something but I forget what <was painted by Bao>.' (Potsdam 2007:585)

Our data reveal the same flexibility in Tagalog, where a mismatch between the Philippine-type actor voice and patient voice is possible under sluicing (6).

(6) *Voice mismatch in Tagalog sluicing: Actor voice - Patient voice*
Bumili si Maria ng bulaklak pero hindi ko alam kung anong klaseng
buy.AV PIVOT Maria INDF.ACC flower but NEG 1SG know if what kind
bulaklak <ang binili ni Maria>.
flower PIVOT buy.PV GEN Maria
'Maria bought a flower but I don't know what kind of flower <was bought by Maria>.'

Whether or not voice mismatch under clausal ellipsis is possible seems to distinguish

Philippine-type voice from English-type voice. In this paper, we take this difference to be real and use it to probe the voice system of Javanese. We first show that Javanese free exceptives, as in (7), are instances of clausal ellipsis, and that Javanese sluicing and free exceptive constructions allow voice mismatch with their antecedents (8). We then extend our investigation to other Western Austronesian languages.

(7) *Javanese clausal free exceptive*
Kabeh arek cilik iku nangis, kecuali Hasan ~~sing gak nangis~~.
all youngster little DEM AV.cry except Hasan ~~REL NEG AV.cry~~
'All the children are crying, except Hasan (is not crying).'

(8) *Voice mismatch allowed in Javanese free exceptives: Active-Passive*
Arek-arek iku nyolong kabeh permen-e kecuali permen kojek <sing gak
child-RED DEM AV.steal all sweet-DEF except sweet lollipop <REL NEG
di-colong ambek arek-arek>.
PASS-steal by child-RED>
'The children stole all the sweets except lollipops (weren't stolen by the children).'

The facts outlined above yield two major conclusions: First, to account for the contrasting behaviour between English-type and Philippine-type voice systems, Indo-European voice and Austronesian voice should not be analyzed in the same way. Second, given the acceptability of voice mismatch in Javanese, the Javanese voice system is better analyzed as Philippine-type.

The layout of the paper is as follows: Section 2 pursues Javanese free exceptives, arguing that they are a clausal ellipsis construction. We then show in more detail that voice mismatches are possible in Javanese free exceptives and sluicing. Section 3 discusses the theoretical background for analyzing the (im)possibility of voice mismatch under clausal ellipsis in English vs. Javanese and accounts for the difference. Section 4 brings in data from three other Austronesian languages, Indonesian, Acehnese, and Puyuma, all of which have been described as having an English-type passive in addition to a Philippine-type voice system. The expectation is that voice mismatch will then not be allowed with these passives. Indonesian confirms the prediction; however, Acehnese and Puyuma mysteriously do not. Section 5 summarises our findings.

Except where otherwise indicated, the data presented in this paper come from primary fieldwork on East Javanese, Pidie Acehnese, Standard Indonesian, Manila Tagalog, and Nanwang Puyuma, through elicitation and acceptability judgments from native speakers over the period of 2022 to 2024.

2. Clausal Exceptives in Javanese

2.1. Javanese syntax basics

Javanese (ISO 639-2 *jav*), a Malayo-Polynesian language spoken on the island of Java, Indonesia, is commonly described as an Indonesian-type language with a three-way voice system: actor voice (9a), object voice (9b), and the often contested "passive" voice marked by the verbal prefix *di-* (9c), which we neutrally refer to as the *di*-construction throughout this paper.

(9) *The three-way voice system in Javanese*

a. Hasan ng-irim hadiah. — Actor Voice (AV)
Hasan AV-send gift
'Hasan sent a gift.'

b. Hadiah iki tak/mbok=kirim. — Object Voice (OV)
gift this 1SG/2SG=OV.send
'I/you sent this gift.'

c. Hadiah-e di=kirim ((ambek) Hasan). — *di*-construction
hadiah-DEF DI/3=send by Hasan
'The gift was sent (by Hasan).'

The actor voice (9a) is marked by a homorganic nasal prefix on the verb and S(AUX)VO word order. The object voice (9b) is characterized by a bare verb, with a strict person and adjacency constraint imposed on the agent: the agent must be first or second-person singular and must surface immediately before the verb. We assume that they are clitics: *tak*= '1SG' and *mbok*= '2SG'. The theme can appear pre-verbally or remain in a post-verbal position. Lastly, the so-called passive voice (9c) is characterized by a clitic *di*=, traditionally glossed as a passive marker. The agent is restricted to the 3rd person. The theme can either appear pre-verbally or remain post-verbal. There is a debate in the literature regarding the *di*-construction, whether it is a true passive or an instance of the object voice restricted to the third person (e.g. Wedhawati 2006; Robson 2014; Kraußе 2017; Patrianto & Chen 2023). Given its behaviour in clausal ellipsis, we will argue that *di*= is a third-person restricted OV marker. For the sake of simplicity, we use the term 'pivot' to refer to the syntactically pivotal phrase eligible to surface in the preverbal position in each voice.

Like in many other western Austronesian languages, Ā-extraction in Javanese is restricted by a pivot-only constraint (Keenan & Comrie 1977): only the syntactically pivotal phrase in a given clause may extract. Nonpivots cannot be directly questioned. This is illustrated in (10). Questioning the agent pivot in AV is grammatical, (10a), but questioning the nonpivot theme directly is not, (10b). Instead, the voice system must first be used to make the theme the pivot via the use of OV, at which point it can be questioned, (10c).

(10) *The pivot-only constraint in Javanese Ā-extraction*

a. Sopo (sing) mangan sego?
WH REL AV.eat rice
'Who ate rice?'

b. *Opo sing Joko mangan?
WH REL J AV.eat
(Intended: 'What did Joko eat?')

c. Opo *(sing) ta'-pangan?
WH REL 1SG-eat.OV
'What did I eat?'

Javanese employs the exceptive marker *kejaba* (native) or *kecuali* (a loan from Indonesian) for both connected and free exceptives.[2] Connected exceptives (11a) are those in which the excep-

[2] *Kejaba* has fallen out of use with younger generations and although recognised as meaning 'except', our Javanese

tive phrase is adjacent to and forms a constituent with its associate. In free exceptives (11b), the exceptive phrase is not adjacent to the associate, nor does it form a constituent with it. We will only be concerned with free exceptives in this paper.

(11)a. Kabeh arek kecuali Joko ngejar asu iku. Connected exceptive
all child except Joko AV.chase dog DEM.DIST
'All the children except Joko chased the dog.'

b. Kabeh arek ngejar asu iku kecuali Joko. Free exceptive
all child AV.chase dog DEM.DIST except Joko
'All the children chased the dog except Joko.'

Recent work has argued that free exceptives in a wide variety of languages are clausal ellipsis constructions (Pérez-Jiménez & Moreno-Quibén 2012; Soltan 2016; Potsdam & Polinsky 2019; Vostrikova 2019, 2021; Potsdam 2022; see Harris 1982 and Merchant 2001 for this idea). In the next section, we support this view for Javanese.

2.2. Javanese exceptives as a clausal ellipsis construction

Four diagnostics indicate that Javanese free exceptives are underlyingly clausal. First, the elided clausal material can be pronounced (2.2.1); second, non-DP constituents can be exceptions (2.2.2); third, a free exceptive yields ambiguity in sluicing (2.2.3); and finally, multiple exceptions are allowed (2.2.4).[3]

2.2.1. Full expression of a clause

A straightforward diagnostic to determine whether a free exceptive is clausal is whether a full clause may be pronounced. Example (12) shows that a full clause can be pronounced in a Javanese free exceptive (although considered redundant).

(12) Kabeh arek cilik iku ngguyu, kecuali Hasan (sing gak ngguyu).
all child little DEM.DIST AV.laugh, except Hasan REL NEG AV.laugh
'All the children laugh except Hasan (is the one who does not laugh).'

Notably, the full clause must be in the form of a cleft, as in (12). This plausibly arises because the exception is the new, focused material and a cleft is the focus construction in Javanese. The use of a non-cleft yields a distinct reading where *kecuali* is interpreted as 'unless':

(13) Kabeh arek cilik iku ngguyu, kecuali Hasan gak ngguyu.
all child little DEM.DIST AV.laugh, unless Hasan NEG AV.laugh
'All the children will laugh unless Hasan laughs.'

language consultants (East Java; under 40) preferred using the Indonesian loanword, *kecuali*. Using either *kecuali* or *kejaba* did not affect acceptability judgments.

[3] See Potsdam 2024 for details. For reasons of space, we do not present two further arguments: a clausal adverb can modify the exception and the exception is island sensitive.

The voice of the clause is restricted by the pivot-only extraction restriction. The exception, *Hasan* in (12), is fronted through focus movement and thus must be the pivot of the missing clause. Given that a full clause following the exceptive marker is possible, ellipsis easily derives the free exceptive with the correct clausal interpretation.

2.2.2. Non-DP exceptives

Exceptions in clausal exceptives can be non-nominal, whereas those in a phrasal exceptive must be nominal. The possibility of non-DP exceptions follows if the mechanism that allows the exception to avoid ellipsis is insensitive to the category of the exception. If the exceptive marker directly selected the exception, however, the exception would likely be restricted to being a DP. In Javanese, PP exceptions in free exceptives are acceptable, but redundant (14a). This contrasts with connected exceptives (14b), which do not have a hidden clausal structure and where the inclusion of a preposition results in ungrammaticality.

(14)a. Aku muteli kembang nang saben kebon saksuwene mongso ketigo kejaba (nang)
1SG AV.pick flower PREP every park during season dry except PREP
kebon iki.
park DEM.PROX
'I pick flowers in every park during the summer, except in this park.'

b. Aku muteli kembang nang saben kebon kejaba (*nang) kebon iki
1SG AV.pick flower PREP every park except PREP park DEM.PROX
saksuwene mongso ketigo.
during season dry
'I pick flowers in every park except (*in) this park during the summer.'

2.2.3. Ambiguity in sluicing

Stockwell and Wong (2020) propose a diagnostic based on sluicing interpretations to detect clausal exceptives (initially noted in Merchant 2001:22). The premise is that a clausal exceptive should provide an additional clause to serve as the antecedent to a sluiced clause. Consider (15). The sluiced clause introduced by *why* can be interpreted in two ways. The antecedent can be the matrix clause as in the interpretation in (15a), or it can be the exceptive clause as in (15b). This second interpretation requires that there be a hidden clausal structure in the exceptive to serve as an antecedent. A non-clausal exceptive, such as the connected exceptive in (16), does not have this second meaning in (16b).

(15) Nobody liked the musical except Mary but I don't know why.
(a) ...but I don't know why <nobody liked the musical except Mary>
(b) ...but I don't know why <Mary liked the musical>

(16) Nobody except Mary liked the musical but I don't know why.
(a) ... but I don't know why <nobody except Mary liked the musical>
(b) *... but I don't know why <Mary liked the musical>

The same interpretations are seen in Javanese free and connected exceptives. In (17), the free exceptive, the sluiced clause can take as its antecedent the main clause (17a) or only the clause that has been elided in the exceptive (17b). The second meaning requires that there be a clause in the exceptive phrase. This reading is once again unavailable in the phrasal connected exceptive, (18).

(17) Kabeh jajan di=pangan ambek de'e, kecuali kismis tapi aku gak ngerti opo'o.
all snack DI=eat by 3SG, except raisins but I NEG AV.know why
'The snacks were eaten by him, except raisins, but I don't know why.'
(a) ... but I don't know why <the snacks except raisins were eaten by him>
(b) ... but I don't know why <raisins weren't eaten by him>

(18) Kabeh jajan kecuali kismis, di=pangan ambek de'e tapi aku gak ngerti opo'o.
all snack except raisins DI=eat by 3SG but I NEG AV.know why
'All the snacks except raisins were eaten by him, but I don't know why.'
(a) ... but I don't know why <the snacks except raisins were eaten by him>
(b) *... but I don't know why <raisins weren't eaten by him>

2.2.4. Multiple exceptions

Free exceptives in both English and Javanese can host multiple exceptions, each related to a different quantificational associate in the main clause (19). This is argued to be possible only if the exceptive has a clausal structure. See Moltmann (1995), Pérez-Jiménez & Moreno-Quibén (2012), and Potsdam (2024) for details of the argument.

(19) Saben arek lanang tak=kenal-no nyang saben arek wedok kecuali Hasan
every child male 1SG=OV.know-APPL PREP every child female except Hasan
nyang Ema.
PREP Ema
'I introduced every boy to every girl except Hasan to Ema <I did not introduce>.'

2.3. Voice mismatch in Javanese exceptives and sluicing

The diagnostics above indicate that free exceptives in Javanese are clausal ellipsis constructions; the missing clause is a cleft, with the exception serving as the focus of the cleft. Because of the language's pivot-only restriction on Ā-extraction, the focus is the pivot of the clefted clause, which determines the voice of the elided clause. Because the voice of the elided clause is known, Javanese is an ideal testing ground for examining the possibility of voice mismatch in clausal ellipsis constructions. This section shows that, to the extent that we can test, voice mismatches across all three voices are possible in both free exceptives and sluicing.

We start with three voice mismatch combinations in free exceptives in (20). They are all acceptable, with or without the material in parentheses. The voice on the verb in the elided clause

is that required by the pivot-only restriction on Ā-movement. Most other combinations result in voice matching or are impossible to construct due to incompatible person restrictions on the agent in OV (1st/2nd person only) and the *di*-construction (3rd person only).

(20) *Voice mismatches in Javanese clausal ellipsis*

a. *AV–Di-construction*

Asu-ku nyokot wong-wong (kabeh) kecuali Ema (sing gak di=cokot ambek dog-1SG.POSS AV.bite person-RED all except Ema REL NEG DI=bite PREP asu-ku).

dog-1SG.POSS

'My dog bit all the people except Ema (did not get bitten by my dog).'

b. *AV–OV*

Aku tuku kabeh buah-buahan pas mongso ketigo kecuali apel (sing gak 1SG AV.buy all fruit-RED when season dry except apple REL NEG tak=tuku).

1SG=OV.BUY

'I bought all the fruit during the summer except apples (were not bought by me).'

c. *Di-construction–AV*

Gedang-e di=pangan (ambek) arek-arek iku kabeh kecuali Joko (sing gak banana-DEF DI=eat PREP child-RED DET all except Joko REL NEG mangan gedang-e).

AV.eat banana-DEF.

'The bananas were eaten by all the children except Joko (did not eat the bananas).'

Voice mismatch is also possible in sluicing (21), where again the optionally elided clause is subject to the pivot-only extraction restriction.

(21) *Voice mismatches in Javanese sluicing*

a. *AV–Di-construction*

Hasan njiwet wong tapi aku gak eruh sopo (sing di=jiwit (ambek) Hasan).
Hasan AV.pinch person but 1SG NEG know who REL DI=pinch by Hasan

'Hasan pinched a person but I don't know who (was pinched by Hasan).'

b. *Di-construction–AV*

Apel-e Eva di=pangan wong (liyo), tapi de'e gak eroh sopo (sing apple-DEF.POSS Eva DI=eat person (other), but 3SG NEG know who REL mangan apel-e).

AV.eat apple-POSS

'Eva's apple was eaten by some person but she doesn't know who (ate her apple).'

c. *AV–OV*

Aku	mecah-ne	barang	tapi	aku	gak	ileng	opo	(sing
1SG	AV.break-CAUS	thing	but	1SG	NEG	remember	what	REL

tak=pecah-ne).

1SG=OV.break-CAUS

'I broke something but I don't remember what (was broken by me).'

Javanese clausal ellipsis constructions allow voice mismatch between the antecedent and the missing clause. This differs from English which does not allow such a mismatch. Our first conclusion is that the Javanese voice system is in essence Philippine-type, as in Malagasy, despite the language's typological similarities with Indonesian-type languages such as Indonesian and Acehnese. This conclusion is in line with a recent investigation of Javanese voice, which has concluded based on independent evidence that Javanese voice is best analyzed as Philippine-type, where voice alternation indexes obligatory topicalization (Patrianto & Chen 2023). In the next section, we turn to a theoretical account of this difference between English-type and Philippine-type voice systems.

3. The (im)possibility of voice mismatch: English vs. Javanese

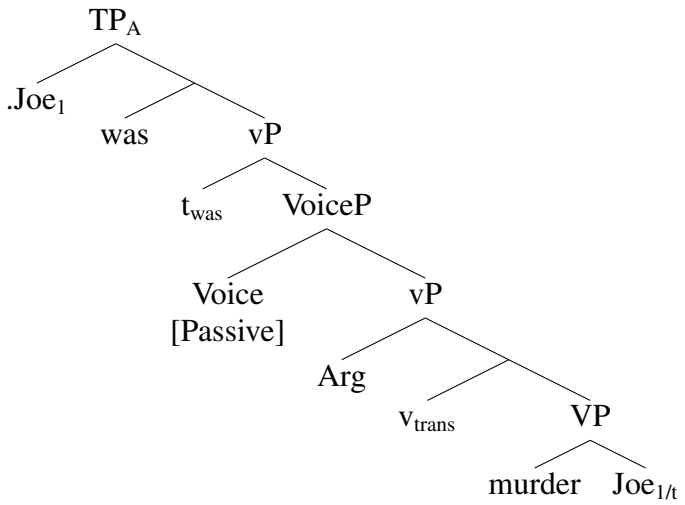
To address the contrast between English and Javanese, we begin with a discussion of our assumptions regarding ellipsis licensing. We then sketch our analyses of English voice and Javanese voice.

It has been widely accepted in the literature that there is an identity condition on ellipsis: there must be some kind of identity between the antecedent and the elided material in order for ellipsis to succeed. Merchant (2001) proposes a purely semantic condition in which the antecedent and the elided must entail each other (to first approximation). Such a condition fails to account for the impossibility of voice mismatch in English because passive and active clauses are truth-conditionally equivalent and entail each other. In response, much recent work has acknowledged that there must be a syntactic component to the identity condition. The recent literature has argued for hybrid syntactic/semantic conditions as well as purely syntactic conditions (Merchant 2013; Chung 2006, 2013; AnderBois 2011, 2014; Tanaka 2011; Barros 2014; Weir 2014; Rudin 2019; Ranero 2021).

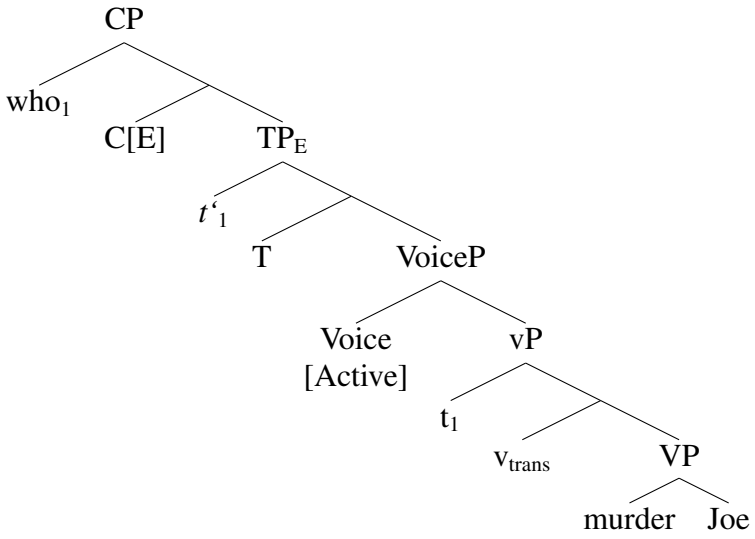
Merchant (2013) argues that voice mismatch in clausal ellipsis constructions is disallowed in English-type languages, an instance of a more general prohibition on argument structure mismatches in clausal ellipsis (Chung, Ladusaw, & McCloskey 1995; Merchant 2001, 2013; Chung 2013; Rudin 2019; others). His analysis of the ungrammatical English voice mismatch in (22) is below. The structure of the passive antecedent is (22a), while that of the active elided clause is (22b). The two clauses mismatch in their featural specification on Voice, which is internal to the ellipsis site. [Passive] does not match [Active], resulting in an inability to elide.

(22) *Joe was murdered (by someone) but we don't know who. (Merchant 2013:92)

a. Passive Antecedent



b. Active Ellipsis Target



Assuming that the same syntactic identity condition holds for Javanese, the acceptability of voice mismatch in Javanese indicates that voice in English and that in Javanese should not be analyzed in the same way. We argue that voice mismatches are allowed if the voice system is not an argument structure alternation and is located at least partly external to the ellipsis site. This provides an account for the possibility of voice mismatch.

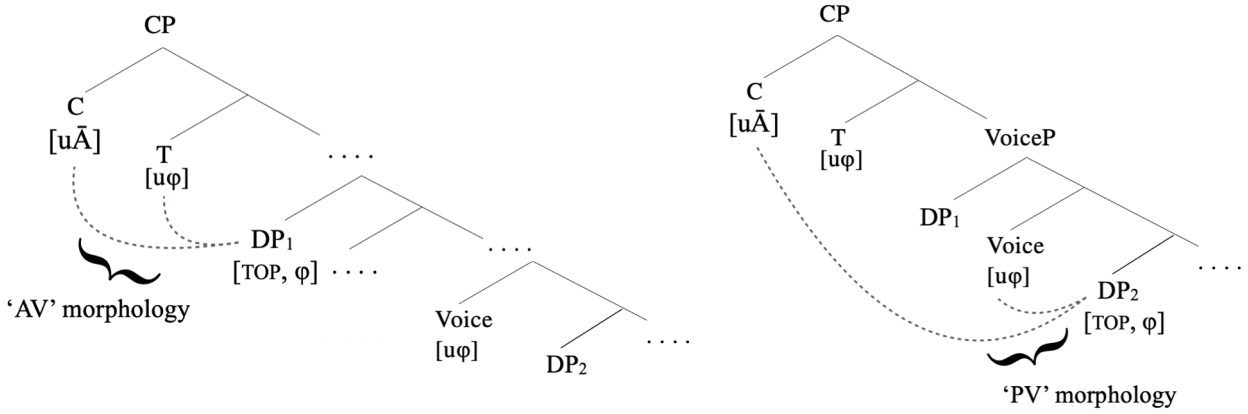
This conclusion aligns with recent Ā-approaches to Philippine-type voice, which, despite minor differences among authors, maintain that some of the relevant voice-related features in Philippine-type Austronesian languages, such as Tagalog and Malagasy, are encoded in the left periphery, external to VoiceP and the ellipsis site. In this view, Philippine-type voice is fundamentally different from Indo-European-type voice and should not be considered a variant of the latter, despite the use of the same terminology (e.g. Chung 1994; Pearson 2001, 2005; Chen 2017, 2021; see also Shibatani 1988 and Richards 2000 for a similar claim).

This line of approach is illustrated below with a recent analysis from Chen (2017, 2021). Following Pearson's (2001, 2005) insight into Malagasy voice, this approach positions Philippine-type voice in the left periphery, associated with information structure (specifically, topichood) and independent from the English-type voice alternation (which is hosted within VoiceP). Specifically, voice affixes are analyzed as the morphological realization of parallel chain relations that involve

Agree with [uTOP]: AV morphology is claimed to be the spell-out of the parallel chains of (abstract) topic agreement and (abstract) subject agreement probing the same goal (23a); PV morphology marks the parallel chain formed with the convergence of (abstract) topic agreement and (abstract) object agreement on the same goal (23b). In other words, the pivot phrase in these languages is assumed to be a topic that agrees with [uTOP] on C and Ā-moves to the left periphery (either through overt or covert Ā-movement).[4]

(23) *The Ā-approach to Philippine-type voice in Tagalog, Puyuma, Amis, and Seediq*

a. Actor voice b. Patient voice (Chen 2021)



In line with Pearson's Ā-approach to Malagasy, voice is an Agree relation between C and a topic DP. Looking at the two trees above, voice mismatch is possible if we assume that [TOP] is ignored in computing syntactic identity. It can be on different DPs and not cause a violation of the identity condition. Unlike English-type voice, there is no argument structure alternation or featural differences on the Voice head in Philippine-type voice systems. Thus, Philippine-type languages allow voice mismatch under clausal ellipsis, as shown earlier in Malagasy (5) and Tagalog (6).

In this view, voice mismatches in Javanese are fully expected if it possesses a Philippine-type Ā-oriented voice system. This conclusion is supported by a recent investigation of Javanese voice. Drawing on systematic diagnostics of Javanese's three voices, Patrianto & Chen (2023) argue that the Javanese voice system is indeed underlyingly Philippine-type, featuring topic-like pivots, akin to those in Philippine-type languages such as Tagalog and Malagasy (Richards 2000, Chen 2017; Pearson 2005). The authors argue that voice alternations in Javanese mark obligatory topicalization reflecting Agree with an Ā-feature ([TOP]). It is therefore distinct from English-style voice alternation. Specifically, what is known as the passive in Javanese (the *di*-construction from earlier in this section) is argued to be an instance of nonsubject topicalization that involves no external argument demotion, the same as in the Javanese object voice. In this view, OV and the *di*-construction are collapsed in Javanese as instances of OV distinguished by the person of the subject. Both mark nonsubject topicalization. The possibility of various types of voice mismatch (20)–(21) in Javanese follows from this analysis.

This observation from Javanese suggests that the possibility of voice mismatch may serve

[4] See Richards (2000) for a covert movement analysis for Tagalog pivots and Chen (2017) for a proposal of the presence or absence of overt Ā-movement as a parametric design of Philippine-type languages.

as a tool for identifying the locus of voice in western Austronesian languages and for determining whether an apparent passive (such as the *di*-construction in Javanese) is Philippine-type or English-type.

4. Austronesian languages with English-type voice

Thus far, we have concluded that Javanese has a Philippine-type voice system and that such systems allow voice mismatches under clausal ellipsis because they are not voice in the Indo-European sense of being argument structure alternations but, rather, Ā-agreement.

This analysis predicts that Austronesian languages with an English-type passive will not allow voice mismatch. This section explores that prediction. Several Austronesian languages have previously been described as having an English-type passive. These languages include Indonesian (Fortin 2007; Cole et al. 2008) and Acehnese (Lawler 1977; Asyik 1987; Legate 2012, 2014). In section 4.1 we show that our prediction is borne out for the passive in Indonesian. Sections 4.2 and 4.3, in contrast, explores data from Acehnese and Puyuma and demonstrate that mismatches are mysteriously still allowed with its English-type passive.

4.1. Indonesian

Indonesian (ISO 639-3 *ind*) is a Malayic language also classified as Indonesian-type. Like Javanese, it possesses a voice system commonly described as three-way (e.g. Chung 1976; Guilfoyle et al., 1992; Sneddon 2010): an actor voice construction that promotes the agent/initiator (24a) to pivot status, marked by an optional prefix *meN-*, an object voice construction characterized by an unmarked bare verb (24b), and an English-type passive (24c) marked by the prefix *di-*. This construction allows optional post-verbal prepositional phrases headed by *oleh* 'by' which introduces the agent/initiator. When the PP is present, *oleh* is optional if the agent is present and immediately right adjacent to the verb. This fact (amongst others) has been taken as evidence that *oleh* marks a demoted agent (Guilfoyle et al. 1992) and the construction is an English-type passive.

(24)a. *Actor Voice*
Siti sudah (men)-cuci baju-nya.
Siti already MENG-wash clothes-3SG.POSS
'Siti washed her clothes.' (Fortin 2007:23; glosses ours)

b. *Object Voice*
Baju sudah Siti cuci.
clothes already Siti wash
'Clothes were washed by Siti.' (Fortin 2007:23)

c. *Canonical passive*
Baju sudah di-cuci (oleh) Siti.
clothes already DI-wash by Siti
'Clothes were washed by Siti.' (Fortin 2007:24)

Given this picture, the expectation is that the canonical passive should disallow voice mismatch under clausal ellipsis, whereas the object voice should pattern with Philippine-type voice and allow voice mismatch. These predictions are borne out, as Fortin 2007 observes for sluicing (25).

(25)a. *AV-OV*

Saya tahu Ali membeli sesuatu, tapi (saya) tidak tahu apa (yang dia
1SG know Ali MENG.AV-buy something but 1SG NEG know what REL 3SG
beli).
OV.buy
'I know Ali bought something, but I don't know what (he bought).' (Fortin 2007:163, 164)

b. *Passive-AV*

Dapur itu sudah di-bersihkan, tapi saya tidak tahu siapa *(yang
kitchen DEM already PASS-clean but 1SG NEG know who REL
membersihkan-nya).
MENG.AV.clean-3SG
'The kitchen was cleaned, but I don't know who *(cleaned it).' (Fortin 2007:174, 175)

In (25a), the antecedent clause is in AV and the elided clause is in OV. Ellipsis is optional but allowed. In (25b), the antecedent clause is a passive and the sluiced clause is in AV. Ellipsis is unacceptable here, as in the English translation, and the sluiced clause must be pronounced. As with Javanese, we can ascertain the voice of the elided clause based on the fact that Indonesian obeys a pivot-only restriction on Ā-extraction. The voice of the sluiced clause in (25a) is OV because the wh-phrase is the theme, and the voice of the sluiced clause in (25b) is AV because the wh-phrase is the agent/initiator.[5] Indonesian thus supports the general picture we have presented regarding the difference between English-type and Philippine-type voice.[6] In the next section, we turn to Acehnese. Although it has a voice system seemingly identical to Indonesian, the mismatch options are not the same.

4.2. Acehnese

The second language which we test our prediction against is Acehnese (ISO 639-3 *ace*), a Chamic language spoken by the Acehnese people in Sumatra, Indonesia. Like Javanese and Indonesian, Acehnese also possesses a three-way voice system and has been argued to contain an English-type passive (Legate 2012, 2014; Patrianto & Chen 2023). To all appearances, it looks very similar to Indonesian.

The Acehnese actor voice (26a) has the agent in subject position agreeing with the verb. It instantiates the language's neutral SVO order. The passive (26b) has the theme in subject position. Exceptionally, the verb still agrees with the agent, which is in a post-verbal position introduced by the preposition *le* 'by' (see Legate 2012 for discussion and analysis of this unusual agreement pattern). The resulting word order is the same as in the Indonesian passive. One difference between

[5] The voice of the elided clause in (25a) could be passive; however, the possibility of voice mismatch in (25a) but not (25b) would be mysterious. OV in (25a), as shown, is also an option and allows us to make sense of the data.

[6] Chamorro sluicing also disallows voice mismatch (i) (Chung 2006, 2013). Space considerations prevent us from presenting the data in detail.

(i) *Chamorro*

??Pära ufan-ma-gächa', lao ti in-tingu'.
FUT AGR-PASS-catch but not AGR-know
(Intended: 'They'll be caught, but we don't know who [will catch them].) (Chung 2006:16)

the Indonesian passive and the Acehnese passive is the lack of any overt passive voice marking on the verb in Acehnese.[7] The Acehnese object voice (26c) also parallels Indonesian OV. The theme is the subject, the verb lacks any agreement, and the agent is immediately pre-verbal.

(26)a. *Actor Voice*
Dokto ka geu-peu-ubat Ibrahim.
Doctor PFV 3POL-CAUS-medicine Ibrahim
'The doctor has treated Ibrahim.' (Legate 2014:47)

b. *Passive*
Ibrahim ka geu-peu-ubat le dokto.
Ibrahim PFV 3POL-CAUS-medicine by doctor
'Ibrahim was treated by the doctor.' (Legate 2014:47)

c. *Object Voice*
Ibrahim ka dokto (*geu)-peu-ubat.
Ibrahim PFV doctor 3POL-CAUS-medicine
'Ibrahim was treated by the doctor.' (Legate 2014:47)

Similar to Indonesian, we predict that voice mismatch under clausal ellipsis will be possible with OV but not with the passive. Free exceptive and sluicing data are below.[8] What we will see is that all voice combinations are acceptable, even with the passive construction.

The data below show that both the passive and OV in Acehnese allow voice mismatch[9]. In (27a), AV and OV can mismatch, not surprisingly, assuming that OV is a Philippine-type voice. As in Javanese and Indonesian, the voice of the elided clause is restricted by the pivot-only extraction restriction (Asyik 1987; Durie 1985).[10] More interestingly, in (27b), a passive clause can serve as the antecedent of an elided AV clause.

(27)a. *Voice mismatch in Acehnese clausal exceptive ellipsis: AV-OV*
Dokto nyan ka geu-peu-ubat banmandum aneuk miet kecuali si Ampon
doctor DEM PFV 3POL-CAUS-medicine all child small except si Ampon
<yang hana dokto peu-ubat>.
REL NEG doctor CAUS-medicine
'The doctor treated all the children except Ampon <was not treated by the doctor>.'

[7] A *di-* prefix is present in Acehnese, as a variant of ji- (Asyik 1987:212), yet it functions as an agreement morphology with a 3sg familiar agent. For example, in (26a) or (26b), if Ibrahim were being treated by a familiar friend instead of a doctor, the agreement morphology would be *di-* instead of *geu-*.

[8] For reasons of space, we assert, without presenting the argumentation, that Acehnese free exceptives are clausal ellipsis constructions, as in Javanese. Like Javanese, Acehnese also uses the Indonesian loanword *kecuali* as its exceptive marker.

[9] It is possible to test the full range of voice mismatch combinations in Acehnese since there are no person restrictions on OV and passive as was seen in Javanese. Space considerations prevent us from presenting all of the data.

[10] As in Indonesian, it cannot be ruled out that the elided clause is passive; however, the ability to pronounce an OV clause strongly suggests that OV is an option.

b. *Voice mismatch in Acehnese clausal exceptive ellipsis: Passive-AV*
Lon ka di-let le mandum asèe kecuali asèe dron <yang hana
1SG PFV 3FAM-chase by all dog except dog 2.SG.POL REL NEG
di-let lon>.
3FAM-chase 1SG
'I was chased by all the dogs except your dog <did not chase me>.'

Voice mismatch is also allowed under sluicing. Consider (28a–b), which demonstrate the possibility of voice mismatch with the Austronesian-type voice contrast (AV-OV) (29a) and a mismatch with the English type passive (passive-AV) (29b).

(28)a. *Mismatch of Austronesian-type voices in Acehnese sluicing: AV-OV*
Guree lon na geu-yue sipeu peu tapi hana deuh lon
teacher 1SG.POSS EXIS 3POL-request.AV something what but NEG audible 1SG
deunge peutra nyan <yang gopnyan yue>.
listen what DET REL 3SG.POL OV.request
'My teacher requested something but I couldn't hear what <was requested by them>.'

b. *Mismatch of English-type voices in Acehnese sluicing: Passive-active*
Pisang keuneuleuh ka ji-bloe le si-droe ureung, tapi hana lon teusoe
banana last PFV 3FAM-buy.PASS by si-self person but NEG 1SG know
soe <yang bloe>.
who COMP buy.AV
'The last banana was bought by someone but I don't know who <bought it>.'

We do not yet have an explanation for why the passive in Acehnese can mismatch with the active AV (27b). Analyses of the Acehnese passive (Legate 2012, 2014) suggest that it has the same behavior and structure as its English counterpart, with the exception that the passive verb agrees with the agent, which we do not think is relevant here.

4.3. Puyuma

A similar puzzle arises from Puyuma (ISO 639-3: *pyu*), where voice mismatches in clausal ellipsis and sluicing are also attested.

Puyuma is a Formosan language spoken in southeastern Taiwan. It possesses a typical Philippine-type voice system with four-way voice alternation. On top of Austronesian voice, it also exhibits two passive-like constructions: an anti-agentive passive marked by the affix *u-* (henceforth Pass_1) and an adversative passive marked by the affix *ki-* (henceforth Pass_2). See details of these two passives in Teng 2020 and Chen 2023. What is important for our current investigation is that both passives involve an unambiguous valency-decreasing operation that eliminates the agent/external argument as a core argument. Therefore, clear evidence shows that both passives are hosted within the core verbal domain (VoiceP), akin to English-type voice.

According to primary fieldwork, voice mismatch is possible in both sluicing and clausal free exceptive constructions (29)–(32). Crucially, the mismatch is acceptable with different com-

binations of Philippine-type voices as well as that of English-style voices (i.e. active voice with either type of passive).

(29) *Mismatch of Austronesian-type voices in Puyuma sluicing: Locative voice–Actor voice*
Tu=trakaw-ay i Sawagu i, ma-ulit=ku dra
3.GEN=steal-LV.(ACT) SG.PIVOT Sawagu TOP, AV-don't.know=1SG.PIVOT C
i manay <na tr<em>akaw>.
SG.PIVOT who PIVOT <AV.(ACT)>steal
'Someone stole from Sawagu, but I don't know who (was the one that stole from him).'

(30) *Mismatch of English-type voices in Puyuma sluicing: Pass$_1$-Active*
M-u-asalr na barasa i, ma-ulit=ku dra
AV-PASS$_1$-move DEF.PIVOT stone TOP AV.ACT-don't.know=1SG.PIVOT C
i manay <na em-asalr>.
SG.PIVOT who PIVOT AV-move
'The stone got moved, but I don't know who (was the one that moved it).'

(31) *English-type voice mismatch in Puyuma clausal exceptive ellipsis: Pass$_2$ + Active*
Ki-pulang=ku dra trawtraw adaman, maumau i Isaw
AV.PASS$_2$-help=1SG.PIVOT INDEF.OBL people yesterday except SG.PIVOT Isaw
<na adri pulang kanku>.
PIVOT NEG help.AV.ACT me
'I was helped by everyone yesterday, except Isaw (was the one who did not help me).'

Crucially, our data also indicate the possibility of clausal ellipsis under simultaneous mismatches of Philippine-type and English-type voices. Consider example (32), which demonstrates the mismatches both of active versus passive voice and between two distinct types of Philippine-type voices: locative versus actor voice. This simultaneous mismatch of both types of voices is unsurprising, given that both passive affixes in Puyuma are compatible with Philippine-type voice alternation in non-ellipsis environments, except for the patient voice (see Chen 2023 for an explanation). The central puzzle, therefore, remains the possibility of English-type voice mismatches, as illustrated in (32).

(32) *Philippine-type and English-type voice mismatches in Puyuma clausal exceptive ellipsis*
Tu=pulang-ay na lalak, maumau i Senten <na adri
3.GEN=help-LV.ACT DEF.PIVOT children except SG.PIVOT Senten PIVOT NEG
ki-pulang>.
AV.PASS$_2$-help
'All children got helped, except Senten (was the one who did not get helped.)'

We have yet to ascertain why all types of voice mismatch are possible in Puyuma during ellipsis and sluicing. However, the logic of our voice mismatch diagnostics leads to the preliminary conclusion that both the Philippine-type and passive voices in Puyuma are actually different from English-type voices, which could explain the observed differences in voice mismatch.

4.4. Summary

Based on our finding that Philippine-type voice, as opposed to English-type voice, allows mismatches under clausal ellipsis, we predicted that Austronesian languages with English-type passives should behave like English and disallow voice mismatch with a non-passive clause. This prediction was only partially confirmed with unexpected counterexamples in Acehnese and Puyuma. On the one hand, the English-like passives in Indonesian and Chamorro behave as expected, disallowing clausal ellipsis when there is voice mismatch. On the other hand, English-like passives in Acehnese and Puyuma are apparent counterexamples and freely allow clausal ellipsis with a mismatching voice.

5. Conclusion, implications, and remaining questions

This paper has used the impossibility of voice mismatch under clausal ellipsis (Merchant 2013) to argue that Philippine-type voice systems in Austronesian languages should not be equated with English-type passive voice, as voice mismatches in Philippine-type systems are permitted. Two clausal ellipsis constructions were used to make this point: sluicing and free exceptives. Free exceptives in some languages are a clausal ellipsis construction and are thus valuable for theorizing about the constraints and conditions on clausal ellipsis. Specifically, we used this diagnostic and free exceptives to argue that the voice alternations in Javanese are of the Philippine-type, supporting analyses of the Javanese voice system in Patrianto and Chen (2023, to appear). We proposed that, in languages with Philippine-type voice, voice mismatch is possible under clausal ellipsis because Philippine-type voice should be analyzed as a type of Ā-agreement and not an argument structure alternation at all.

We document, however, that some Austronesian languages analyzed as having an English-type voice construction, namely, Acehnese and Puyuma, unexpectedly allow voice mismatch in clausal ellipsis constructions. We do not yet have an explanation for this observation.

References

AnderBois, Scott. 2011. *Issues and alternatives.* University of California, Santa Cruz, Doctoral dissertation.

AnderBois, Scott. 2014. The semantics of sluicing: beyond truth conditions. *Language* 90(4), 887–926.

Asyik, Abdul Gani. 1987. *A contextual grammar of Acehnese sentences.* University of Michigan, Doctoral dissertation.

Barros, Matthew. 2014. *Sluicing and identity in ellipsis.* Rutgers, The State University of New Jersey-New Brunswick, Doctoral dissertation.

Chen, Victoria. 2017. *A reexamination of the Philippine-type voice system and its implications for Austronesian primary-level subgrouping.* University of Hawai‘i at Mānoa, Doctoral dissertation.

Chen, Victoria. 2021. Tagalog voice as four bundles of Agree relations: Insights from binding. *Proceedings of the 38th West Coast Conference on Formal Linguistics (WCCFL 38)*, 94–104. Somerville, MA: Cascadilla Proceedings Project.

Chen, Victoria. 2023. When Philippine-type voice meets Indo-European-type voice: Insights from Puyuma. *Journal of East Asian Linguistics* 31, 517–550.

Chung, Sandra. 1976. On the Subject of Two Passives in Indonesian. In C.N. Li (ed.), *Subject and Topic*. Academic Press, New York, 57–99.

Chung, S., 1994. Wh-Agreement and referentiality in Chamorro. *Linguistic Inquiry* 25(1), 1—44.

Chung, Sandra. 2006. Sluicing and the lexicon: The point of no return. In Rebecca T. Cover and Yuni Kim (eds.), *Proceedings of the annual meeting of the Berkeley Linguistics Society* 31, 73–91. Berkeley: University of California, Berkeley Linguistics Society.

Chung, Sandra. 2013. Syntactic identity in sluicing: How much and why. *Linguistic Inquiry* 44, 1–44.

Chung, Sandra, William A. Ladusaw, & James McCloskey. 1995. Sluicing and Logical Form. *Natural Language Semantics* 3, 239–282.

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

Durie, Mark. 1985. *A grammar of Acehnese on the basis of a dialect of North Aceh*. Dordrecht: Foris.

Fortin, Catherine. 2007. *Indonesian sluicing and verb phrase ellipsis: Description and explanation.* University of Michigan, Doctoral dissertation.

Guilfoyle, Eithne, Henrietta Hung, & Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. *Natural language and linguistic theory* 10, 375–414.

Harris, Zellig. 1982. *A Grammar of English on Mathematical Principles.* NY: John Wiley & Sons.

Keenan, Edward L., & Comrie, Bernard. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, 63–99.

Kraußе, Daniel. 2017. *A description of Surabayan Javanese: with special reference to its linguistic etiquette.* Goethe University, Master thesis.

Lawler, John M. 1977. A agrees with B in Achenese: A problem for Relational Grammar. In *Syntax and Semantics 8: Grammatical Relations*, ed. Peter Cole and Jerold Sadock, 219–248. Academic Press.

Legate, Julie. 2012. Subjects in Acehnese and the nature of the passive. *Language* 88, 495–525.

Legate, Julie. 2014. *Voice and v: Lessons from Achenese*. MIT Press.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis.* Oxford: Oxford University Press.

Merchant, Jason. 2008. An asymmetry in voice mismatches in VP-ellipsis and pseudo-gapping. *Linguistic Inquiry* 39, 169–179.

Merchant, Jason. 2013. Voice and Ellipsis. *Linguistic Inquiry 44*(1), 77–108.

Moltmann, Friederike. 1995. Exception sentences and polyadic quantification.*Linguistics and Philosophy* 18(3), 223–280.

Patrianto, Hero, & Victoria, Chen. 2023. Two sides to the same coin: Reappraising Indonesian-type 'passive' and object voice in Javanese. *Proceedings of AFLA* 29, 59–74

Patrianto, Hero, & Victoria, Chen. To appear. Rethinking Indonesian-type passives: Lessons from Javanese. *Proceedings of NELS 53*.

Pearson, Matt. 2001. *The clause structure of Malagasy: A minimalist approach.* University of California, Los Angeles, Doctoral dissertation.

Pearson, Matt. 2005. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory* 23, 381–457.

Pérez-Jiménez, Isabel & Norberto Moreno-Quiben. 2012. On the syntax of exceptions. Evidence from Spanish. *Lingua* 122(6), 582—607.

Potsdam, Eric. 2007. Malagasy sluicing and its consequences for the identity requirement on ellipsis. *Natural Language & Linguistic Theory* 25, 577–613.

Potsdam, Eric. 2018. Exceptives, clausal ellipsis, and island (in)sensitivity. Invited talk, *25th Meeting of the AFLA*: Academia Sinica, Taipei, Taiwan.

Potsdam, Eric. 2022. What Sluices in Malagasy Sluicing? *Oceanic Linguistics* 61(2), 757–790.

Potsdam, Eric. 2024. Introduction to exceptives questionnaire. Ms. University of Florida. https://exceptives.lin.ufl.edu/exceptives-questionnaire/. Downloaded April 1, 2024.

Potsdam, Eric & Maria Polinsky. 2019. Clausal and phrasal exceptives. *Conference Presentation at Generative Linguistics in the Old World (GLOW)* 42.

Ranero, Rodrigo. 2021. *Identity conditions on ellipsis.* University of Maryland at College Park, Doctoral dissertation.

Richards, Norvin. 2000. Another look at Tagalog subjects. In Paul, I. et al. (eds.), *Formal issues in Austronesian linguistics*, 105–116. Kluwer.

Robson, Stuart. 2014. *Javanese grammar for students*. Monash University Publishing.

Rudin, Deniz. 2019. Head-based syntactic licensing in sluicing. *Linguistic Inquiry* 50(2), 253–283.

Shibatani, Masayoshi. 1988. Voice in Philippine languages. In M. Shibatani (ed.), *Passive and voice*, 85–142. Amsterdam: John Benjamins.

Sneddon, James. 2010. *Indonesian: A comprehensive grammar*. New York: Routledge.

Soltan, Usama. 2016. On the syntax of exceptive constructions in Egyptian Arabic. In Stuart Davis & Usama Soltan (eds.), *Perspectives on Arabic Linguistics xvii*, 35–57.

Stockwell, Richard, & Deborah J.M. Wong. 2020. Sprouting and the structure of except-phrases. In Mariam Asatryan, Yixiao Song, and Ayana Whitmal (eds.), *Proceedings of NELS* 50(3), 169–182.

Tanaka, Hidekazu. 2011. Voice Mismatch and Syntactic Identity. *Linguistic Inquiry* 42(3), 470–490.

Teng, Stacy. 2020. The three agent demoting prefixes (ki-, mu-, kur-) in Katripul Puyuma: Their origins and possible development. *Concentric* 46(1), 21—65.

Vostrikova, Ekaterina. 2019. Compositional analysis for clausal exceptives. *Semantics and Linguistic Theory (SALT)* 29, 420–440.

Vostrikova, Ekaterina. 2021. Conditional Analysis of Clausal Exceptives. *Natural Language Semantics* 29, 159–227.

Wedhawati, Sjamsul Arifin. 2006. *Tata Bahasa Jawa Mutakhir*. Penerbit Kanisius: Yogyakarta.

Weir, Andrew. 2014. *Fragments and clausal ellipsis.* University of Massachusetts at Amherst, Doctoral dissertation.