



Asia-Pacific Linguistics

Open Access

College of Asia and the Pacific
The Australian National University

AFLA 22

**The Proceedings of the 22nd Meeting of the
Austronesian Formal Linguistics Association**

edited by

Henrison Hsieh



AFLA 22: The Proceedings of the 22nd Meeting of the Austronesian Formal Linguistics Association

edited by
Henrison Hsieh

The Austronesian Formal Linguistics Association (AFLA) is an organization that promotes the study of Austronesian languages from a formal perspective. The 22nd annual meeting of the Austronesian Formal Linguistics Association (AFLA 22) was held at McGill University in Quebec, Canada in May, 2015. Of the 32 papers presented at the conference, 14 are included in this volume, with contributions on languages including Marshallese, Tagalog, Javanese and Kavalan and on topics covering syntax, semantics and phonology.



Asia-Pacific Linguistics

Open Access

**AFLA 22:
The Proceedings of the 22nd Meeting of the
Austronesian Formal Linguistics Association**

edited by
Henrison Hsieh



Asia-Pacific Linguistics

Open Access

EDITORIAL BOARD: Bethwyn Evans (Managing Editor),
I Wayan Arka, Danielle Barth, Don Daniels, Nicholas Evans,
Simon Greenhill, Gwendolyn Hyslop, David Nash, Bill Palmer,
Andrew Pawley, Malcolm Ross, Hannah Sarvasy, Paul Sidwell,
Jane Simpson.

Published by Asia-Pacific Linguistics
College of Asia and the Pacific
The Australian National University
Canberra ACT 2601
Australia

Copyright in this edition is vested with the author(s)
Released under Creative Commons License (Attribution 4.0 International)

First published: 2016

URL: <http://hdl.handle.net/1885/101155>

National Library of Australia Cataloguing-in-Publication entry:

Creator: Henrison Hsieh, editor.

Title: AFLA 22: the proceedings of the 22nd meeting of the Austronesian Formal Linguistics Association / Henrison Hsieh, editor.

ISBN: 9781922185303 (ebook)

Series: Asia-Pacific Linguistics; A-PL 27.

Subjects: Austronesian languages – Congresses.

Other Creators Henrison Hsieh, editor.

Australian National University; Asia-Pacific Linguistics.

Dewey Number: 499.2

Cover: AFLA 22 logo designed by Lauren Eby Clemens and Michael Yoshitaka Erlewine

Contents

<i>Preface</i>	iii
by Henrison Hsieh	
Eli Asikin-Garmager	1
<i>Sasak: Antipassives and the syntactic dimensions of nasal verb variation</i>	
Lena Borise	15
<i>Sluicing in Tagalog: Strategies and implications</i>	
Victoria Chen and Shin Fukuda	31
<i>Three ways to “steal” an element from a CP: Evidence from Formosan</i>	
Helen Jeoung	49
<i>Voice and register differences in Madurese</i>	
Dong-yi Lin	65
<i>Floating universal quantifier as a base-generated head in the VP periphery</i>	
Kuo-Chiao Lin	82
<i>The gradual path to variable reduplication: Kavalan</i>	

Nicholas Longenbaugh and Maria Polinsky	98
<i>The processing of long-distance dependencies in Niuean</i>	
Eric Potsdam and Daniel Edmiston	121
<i>Extraposition in Malagasy</i>	
Catherine Provencher	139
<i>On the Marshallese semi-glides and underspecified vowels</i>	
Norvin Richards	155
<i>Some notes on Tagalog prosody and scrambling</i>	
Nova Starr	173
<i>General number and semantic incorporation in Tagalog</i>	
Nozomi Tanaka, William O’Grady, Kamil Deen, Chae-Eun Kim, Ryoko Hattori, Ivan Paul M. Bondoc, and Jennifer U. Soriano	191
<i>An agent advantage in Tagalog relative clause comprehension</i>	
Jozina Vander Klok	202
<i>Diagnosing VP-ellipsis in Javanese: Evidence for a non-movement and a movement account</i>	
Ekarina Winarto	220
<i>The Indonesian DP</i>	

Preface

The 22nd annual meeting of the Austronesian Formal Linguistics Association (AFLA 22) was held on 21-24 May, 2015 at McGill University. The program consisted of 28 presentations in addition to four plenary talks by I Wayan Arka, Ileana Paul, Maria Polinsky, and Joey Sabbagh. This volume includes 14 papers from the conference. Abstracts for these papers underwent peer review prior to acceptance to and presentation at the conference. In light of this, the papers included in this volume did not undergo further peer review.

AFLA 22 was partially funded by the McGill Dean of Arts Development Fund, the McGill Department of Linguistics Cremona Fund, Fonds de recherche du Québec – Société et culture (FRQSC) Team Grant 2012-SE-144646, and Social Sciences and Humanities Research Council (SSHRC) Insight Grant 435-2012-0882. The organizing committee consisted of Lauren Eby Clemens, Michael Yoshitaka Erlewine, Henrison Hsieh, Jiajia Su, and Lisa Travis.

We would like to especially thank the undergraduate volunteers from the Society of Linguistics Undergrads at McGill (SLUM) who helped greatly in ensuring the smooth running of the conference. We would also like to acknowledge Lauren Eby Clemens and Michael Yoshitaka Erlewine for designing the conference logo. Finally, we would like to thank Edith Aldridge, I Wayan Arka, Ellen Broselow, Marc Brunelle, Sandra Chung, Peter Cole, William Davies, Daniel Finer, Catherine Fortin, Gabriella Hermon, Edward Keenan, Hilda Koopman, Paul Law, Diane Massam, Hiroki Nomoto, Yuko Otsuka, Bill Palmer, Ileana Paul, Elizabeth Pearce, Matt Pearson, Glynne Piggott, Eric Potsdam, Norvin Richards, Joachim Sabel, Yosuke Sato, Patricia Schneider-Zioga, Hooi Ling Soh, Jozina Vanderklok, Kilu von Prince, Kristine Yu, Elizabeth Zeitoun, and Kie Zuraw for volunteering their time to the review process.

Henrison Hsieh

SASAK: ANTIPASSIVES AND THE SYNTACTIC DIMENSIONS OF NASAL VERB VARIATION*

Eli Asikin-Garmager
The University of Iowa
eli-asikin-garmager@uiowa.edu

Many Austronesian languages spoken in Indonesia have a nasal prefix that occurs on verbs (hence ‘*nasal* verbs’). While certain aspects of the cognate nasal prefix are similar in these languages, Sasak is different because there are various morphologically-distinct nasal prefixes (Austin 2012) and Sasak dialects vary considerably in which nasal prefix(es) they use and how they function. This paper shows how the Sasak prefixes can be described primarily along two syntactic dimensions; whether the verb projects an internal argument and whether an extraction asymmetry is observed. In a Minimalist framework this can be accounted for by assuming parametricity of single features found on little *v*, an idea first proposed by Aldridge (2008) in a phase-based analysis of Indonesian, Malagasy, and Tagalog. Allowing such feature-based parametricity of syntactic heads permits us to account for the syntactic patterns of the Sasak nasal verbs.

1. Background

Sasak dialectal variation in verbal morphology, clitics, and clause structure is traditionally described with reference to one of five ethnolinguistic classifications. These include Ngenó-ngené and Nggetó-nggeté (Eastern Lombok), Menó-mené and Meriaq-meriku (Central and Southern Lombok), and Kutó-kuté (Northern Lombok). Like related languages, Sasak has a system of verbal affixes, including one or two applicative suffixes and multiple passive prefixes, but more unique to Sasak is the fact that multiple nasal prefixes are found to varying degrees in different dialects. In the remainder of this section I will present a basic description of Sasak’s voice system and verbal morphology needed for the later analysis. Section 2 then describes the syntactic patterns observed with Sasak’s nasal prefixes, and section 3 provides a formal generative analysis of one prefix. Section 4 expands this analysis so it accounts for the full range of prefixes, and shows how the data bear on previous phase-based analyses (e.g., Rackowski and Richards 2005, and particularly Aldridge 2008). Finally, section 5 concludes the paper with a brief summary and questions raised by the Sasak data.

Sasak has a three-way voice system that includes actor voice, object voice, and a morphological passive. Actor voice is used here to describe a transitive

* Thanks to William D. Davies, Alice Davison, Paula Kempchinsky, and Peter K. Austin for discussion of data. I am also grateful to Nur Ahmadi for his continued collaboration. Fieldwork in Northern Lombok was supported by *The Stanley - University of Iowa Foundation Support Organization* and The University of Iowa’s graduate student organizations. All errors are my own.

clause in which a preverbal actor occurs with a nasal-prefixed verb, and objects occur post-verbally, as shown in (1). Unlike in Eastern and Northern dialects, the nasal prefix is not obligatorily present with preverbal actors in Central and Southern dialects.¹ The example in (1) also shows an example of a preverbal auxiliary bearing a subject enclitic. When the subject referent is clear from the discourse, speakers often use only the clitic and drop clause-initial DPs.

- (1) Actor voice Ngenó-ngené (E), Pancor (Shibatani, 2008:877)
 loq Ali wah=*ne* ngirim-ang oku surat
 ART Ali PERF=3SG N.send-APPL 1SG letter
 ‘Ali sent me a letter’²

Sasak also has a morphological passive (2), in which a full DP object occurs to the left of a verb bearing the passive prefix (with any auxiliaries optionally intervening between the object and verb), and the actor/agent occurs (optionally) in a by-phrase clause-finally.

- (2) Passive voice Kutó-kuté (N), Sukadana
 a. aku saweq ta-kirim-ang surat (isiq sawa-ng=ku)
 1SG already PASS-send-APPL letter by spouse-LNK=1SGPOSS
 ‘I was already sent a letter by my wife’
 b. saweq=*ku* ta-kirim-ang surat (isiq sawa-ng=ku)
 already=1SG PASS-send-APPL letter by spouse-LNK=1SGPOSS
 ‘I was already sent a letter by my wife’

The examples in (2) show how when a verb is prefixed with the passive morpheme, the clitic that occurs on the preverbal auxiliary (reserved for actors/agents in actor voice), is cross-referenced with the patient/recipient instead. Thus, in passive voice, the patient/recipient clitic occurs on the preverbal element, not the actor clitic, presumably because it occupies the structural position from which clitic co-reference can be established.

¹ Further description of the function of nasal verbs in Central/Southern areas is provided in section 2. Given considerable variation even within these dialect areas, I follow the established tradition of providing a dialect and village identifier for each example (Austin, 1998, 2000, 2012; Shibatani 2008). Also, for ease of reference, following the dialect name label, I mark examples with reference to the particular dialect area via the abbreviations (E) for Eastern, (N) for Northern, and (C/S) for Central/Southern.

² *Abbreviations used in the glosses include:*

ACC – accusative; ACT – active; AP – antipassive; APPL – applicative; ART – article; DEF – definite; ERG – ergative; INTR – intransitive; LNK – nasal linker (sometimes when a vowel is followed by a stop consonant across a morpheme or word boundary, a homorganic nasal is epenthesized between the two segments); LOC – locative preposition; N – Sasak nasal prefix (also N¹, N², N³, N⁴) clarified further in section 2; PASS – passive; PERF – perfective; POSS – possessive; REL – relativizer; SG – singular; TR – transitive; Ø – zero morpheme

The remainder of section 2 will describe how these prefixes vary primarily along the following two syntactic dimensions: (1) whether the prefixed verb projects an object (internal argument), and (2) whether an extraction asymmetry is observed.

2.1. Dimension #1: Extraction (a)symmetries

N' in Eastern Sasak is the analog of the Indonesian and Balinese nasal prefix, and required on verbs in the presence of overt preverbal (DP) actors. Conversely, bare verbs (i.e., non-prefixed verbs) occur with preverbal patients (as in the transitive object voice construction in (3)). *N'* is required when the external argument is relativized (5a), whereas when the gap corresponds to the patient (i.e., internal argument extraction), the nasal prefix is obligatorily absent (5b). Relativization data is shown below, but the same asymmetrical pattern of prefixation is found for *wh*-movement and topicalization. This pattern is also found in some northern dialect areas (i.e., Gondang) (Austin, personal communication, March 2015).

(5) *N'* in Ngenó-ngené (E), Selong (Austin 2012:35)

a. relativized actor

dengan	[siq	—	<i>m</i> -beli	buku]	inó
man	REL		<i>N'</i> -buy	book	that

‘That man who bought a book’

b. relativized object

buku	[siq	—	<i>mèq</i> =beli]	inó
book	REL		2SG=buy	that

‘That book which you intend to buy’

In contrast to the data shown in (5), in some northern village areas (i.e., Sukadana), the nasal prefix is found on verbs in the presence of both external argument extraction (6a), as well as with internal argument extraction (6b).⁵

(6) *N'* in Kutó-kuté (N), Sukadana

a. relativized actor

amaq	siq	<i>n</i> -talet	jagung	nónó	pintér
man	REL	<i>N'</i> -plant	corn	that	smart

‘The man who planted the corn is smart’

⁵ Most speakers report the sentence in (6b) as ungrammatical without the nasal prefix, but one consultant judged both the prefixed and bare form as acceptable. More data is needed in order to determine whether additional speakers also show optionality and whether or not this may be an effect of register. For example, it is possible that in more colloquial speech such prefixation is more optional.

b. relativized object

jagung siq	<i>n</i> -talet=ku	maiq	gati
corn REL	<i>N</i> ^l -plant=1SG	delicious	very

‘The corn that I planted is very delicious’

Concerning the pattern shown in (6), in their survey of Malay dialects, Cole, Hermon, and Yanti (2008) describe how the presence of a nasal prefix with extracted objects also occurs in Kuching Malay. This contrasts with the extraction asymmetry observed for *N*^l in Eastern Sasak, Indonesian *meN*-, and the Balinese cognate suffix. Moreover, it highlights how this dimension of syntactic variation is not limited to Sasak.

2.2. Dimension #2: Antipassives

Section 2.1 shows how with *N*^l in Eastern dialects and some northern village areas an extraction asymmetry is observed, whereas the prefix is found with both internal and external argument extraction in other northern sub-dialects. This section presents data which show that a separate nasal prefix in Northern Sasak (*N*⁴) functions as an antipassive marker, occurring in intransitive clauses which have only an overt actor.

While both *N*^l and *N*⁴ can combine with many of the same transitive verb bases, in the case of *N*^l, an object is obligatorily present (hence the grammaticality of (7a) but not (7b)). Conversely, in the presence of *N*⁴, the object is obligatorily absent (hence a detransitivized clause), and this accounts for the grammaticality of (8a) but ungrammaticality of (8b).

(7) *N*^l occurs in the obligatory presence of direct objects (Sukadana (N))

- | | | | | | |
|----|--------------|-----------------------------|--------------|-------|------------|
| a. | mèlé-ng=ku | <i>n</i> -séran | <i>burón</i> | kón | pawang=nó |
| | want-LNK=1SG | <i>N</i> ^l -hunt | <i>deer</i> | place | forest=DEF |
- ‘I want to hunt deer in the forest’
- | | | | | |
|----|--------------|-----------------------------|-------|------------|
| b. | *mèlé-ng=ku | <i>n</i> -séran | kón | pawang=nó |
| | want-LNK=1SG | <i>N</i> ^l -hunt | place | forest=DEF |
- ‘I want to hunt deer in the forest’

(8) *N*⁴ occurs in the obligatory absence of overt objects (Sukadana (N))

- | | | | | |
|----|--------------|-----------------------------|-------|------------|
| a. | mèlé-ng=ku | <i>menyéran</i> | kón | pawang=nó |
| | want-LNK=1SG | <i>N</i> ⁴ .hunt | place | forest=DEF |
- ‘I want to hunt in the forest’
- | | | | | | |
|----|--------------|-----------------------------|--------------|-------|------------|
| b. | *mèlé-ng=ku | <i>menyéran</i> | <i>burón</i> | kón | pawang=nó |
| | want-LNK=1SG | <i>N</i> ⁴ .hunt | <i>deer</i> | place | forest=DEF |
- ‘I want to hunt deer in the forest’

Despite various sub-dialectal syntactic differences, there is only one syntactic difference in terms of N^4 in Sukadana and Gondang (both North Lombok). In Sukadana, objects are disallowed with N^4 , but in Gondang they are optionally permissible, but with a non-specific restriction. So, whereas in Sukadana objects are disallowed with N^4 (8b), in Gondang they are subject to a semantic restriction, as shown in (9).

- (9) N^4 occurs with non-specific objects in Gondang (North Lombok)
- | | | | | | | | |
|-------|-----|-----|-----------------|-------|------|-----|--------------|
| araq | tau | ia | <i>mem</i> -bau | buaq | pir | léq | kebòn=nya |
| exist | man | 3SG | N^4 -pick | fruit | pear | LOC | garden=3POSS |
- ‘There is a man picking pears in his garden’
 (Ahmadi 2009, as cited in Austin 2012)

The semantic restriction observed in (9) is strikingly similar to how oblique objects in antipassive constructions must be interpreted as indefinite in Tagalog (Aldridge 2012) and how antipassives are preferred in Chamorro when objects are non-specific (Cooreman 1994).

The use of a nasal prefix in detransitivized clauses is not limited to Northern Sasak dialects. In section 1, it was noted how a nasal prefix often does not occur in actor voice in Central/Southern dialects. This is because the nasal prefix (N^1) in Central/Southern dialects, as well as N^3 in Eastern dialects, have the same detransitivizing effect that we observe for N^4 in Northern dialects (Austin 2012).

2.3. Summary of Sasak’s nasal prefixes’ syntactic patterns

Sections 2.1 and 2.2 present data showing that, despite highly varied morpho-phonological patterns (Austin 2012, Asikin-Garmager 2015), the syntactic variation found with the Sasak nasal prefixes can be described primarily along two syntactic dimensions. These are whether or not an extraction asymmetry is observed and whether or not the prefix occurs in intransitive clauses. This can be found represented in the columns in the table in (10). Dialects have been placed into rows, and recall that superscripts reflect the morphological identities of the prefixes.

- (10) The syntax of Sasak’s nasal prefixes (presentation changed; portions adapted from Austin 2012:44)⁶

Dialect/village	Transitive		Detransitivized
	<i>extraction asymmetry</i>	<i>no extraction asymmetry</i>	
<i>Eastern</i> Ngenó-ngené Nggetó-nggeté	N ¹		N ³
<i>Northern</i> Kutó-kuté, (Gondang area)	N ¹		N ⁴ non-referential objects
Kutó-kuté, (Sukadana area)		N ¹	N ¹ no overt objects
<i>Central/Southern</i> Menó-mené, Menu-meni, Meriatq-meriku	(no N-/Ø alternation observed)		N ¹ non-referential objects

3. Minimalist phase-based approach

The Sasak data presented thus far are relevant for a better understanding of Indonesian voice systems because Indonesian and other related languages have only a single nasal prefix whereas Sasak dialects may have multiple prefixes.⁷ More specifically, the Sasak data bear directly on previous phase-based approaches such as Aldridge (2008) because they provide evidence in favor of the proposal for allowing parametric variation to target single features on syntactic heads. In fact, nearly full parametric variation of a case and EPP feature on little *v* are required in order to account for the Sasak facts. In order to show that this is the case, I will first briefly introduce Aldridge’s (2008) phase-based approach and assumptions, and show how the analysis readily extends to one of the Sasak prefixes. Following, section 5 discusses what accommodations are needed in order to account for the other Sasak prefixes presented in section 2.

⁶ Novel data from Sukadana (Kutó-kuté, North Lombok), along with data from additional sources, have been added in order to provide a fuller picture of the dialects and sub-dialects. Additionally, while I follow Austin’s (2012) use of superscripts, the presentation is reorganized to highlight the syntactic properties of the prefixes. This paper focuses mostly on Eastern and Northern sub-dialects. N² (only found in Eastern dialects) is not included in the table due to lack of additional data concerning its syntactic properties.

⁷ If Balinese /ma-/ is analyzed as a separate form from the actor voice prefix /N-/, then we observe two distinct prefixes in the language. Indeed, Shibatani and Artawa (2003) discuss one of the uses of /ma-/ as an antipassive marker, just as we find for some of the Sasak prefixes. I leave investigation of further comparison of these Sasak and Balinese structures for further work.

Aldridge’s (2008) account of Indonesian, Tagalog, and Malagasy is based on core assumptions from Phase Theory, including that sentences undergo phonological and semantic interpretation in ‘chunks’ or phases (including CP, vP, and definite DP), and that movement must happen cyclically through both vP as well as CP (Chomsky 2001). Secondly, according to the Phase Impenetrability Constraint (PIC) as defined in Adger (2003:340), “feature checking under c-command reaches no further than the specifier of any embedded phase.” Crucially, PIC is assumed to be violated in the case of *strong phases*, but not weak phases (the later including unaccusative and passive v) (Chomsky 2001).

In order to account for the Indonesian, Tagalog, and Malagasy facts, Aldridge proposes that parametric variation can target single features of a syntactic head’s feature bundle (namely a Case and EPP feature in this case), and that various morphemes we observe in these languages are the morphological realization of distinct syntactic heads. The Indonesian cognate nasal prefix *meN-* is posited on par with the antipassive functional head (v_{AP}) in Tagalog, except while *meN-* once may have functioned as an antipassive marker (i.e., occurred in objectless clauses), it has since acquired the characteristic that it can select direct objects. The lack of an EPP on v_{Act} (*meN-*) accounts for the well-known Indonesian extraction asymmetry (i.e., *meN-* occurs with extracted external, but not internal, arguments). With no EPP to draw an internal argument to the strong vP phase edge, internal arguments are unavailable for extraction. However, unlike Tagalog v_{AP} , Indonesian *meN-* has acquired a Case feature. So, although we observe an extraction asymmetry, *meN-*-prefixed verbs now project internal arguments. The proposal that parametric variation can target single features on syntactic heads is shown in (11)—note the acquisition of a [*uCase: ACC*] feature on Indonesian’s v_{AP} .

- (11) Ergativity as a continuum and variation of features on little *v*
(Aldridge, 2008:1455)

<u>Ergative (Tagalog)</u>		→	<u>Split ergative/almost accusative (Indonesian)</u>	
v_{AP}	No EPP feature		v_{AP}	No EPP feature
	No case feature			<i>[uCase: ACC]</i>

Recall from section 2 that, like in Indonesian, *N^l* in Eastern Sasak dialects is required for subject, but not object extraction. Thus, with only some minor differences, we can account for these facts with the featural specifications proposed by Aldridge (2008:1453) for Indonesian, reassigned to the Sasak *v* heads in (12). Key aspects of the analysis include the fact that v_{Act} is spelled-out as the nasal prefix (Indonesian *meN-* and Sasak *N^l*), and that v_{Act} is analogous with Tagalog $v_{Intr/AP}$ except that it has acquired a Case feature.

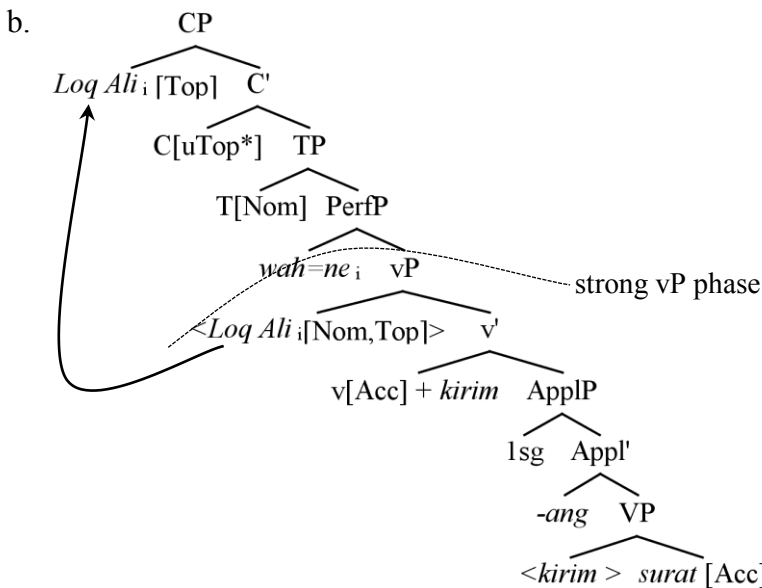
(12) Sasak (N^I) and Indonesian featural specification of heads

head	morpheme	features
v_{Act}	nasal prefix	[uCase:ACC] feature to value with a DP in c-command domain No [uD*] feature
v_{Pass}	passive prefix	No [uD*]; no selection of external argument
v_{Erg}	Ø prefix	Inherent case to assign to the external argument [uD*] feature to draw internal argument to the vP phase edge
T_{Fin}		[uCase: Nom] feature to value with a DP in c-command domain

Referring to the derivation of the Sasak actor voice structure shown in (13b), notice that N^I (little v_{Act}) assigns accusative case to an internal argument, but the syntactic head does not bear an EPP feature. As a result, only the external argument is available for extraction. In this case, this means that only the external argument *loq Ali* is available to check the [uTop*] feature on C. Extraction of the internal argument in the presence of N^I is not possible. The derivation of the illicit construction **oku wah=ne_i ngirim-ang surat isiq loq Ali_i* is elided due to space.

(13) Ngenó-ngené (E), Pancor (actor voice—agent topicalization)

- a. loq Ali_i wah=ne_i ngirim-ang oku surat
 ART Ali PERF=3SG N^I.send-APPL 1SG letter
 ‘Ali sent me a letter’

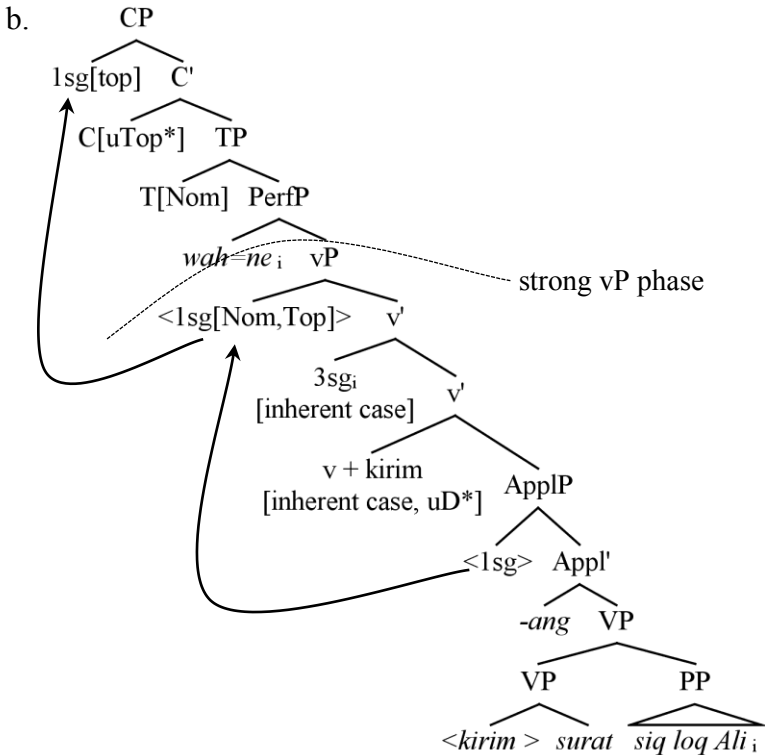


Next, the derivation of an object voice structure in (14b) shows how the featural specification of Sasak v_{Erg} (Ø prefix) (namely the EPP [uD*] feature) will account for the fact that internal arguments can be extracted in the absence of N^I in Eastern Sasak. In (14b) the 1SG pronoun merges in the specifier of the applicative phrase and is assigned accusative case by v_{Act} . Next, v_{Erg} assigns case to the

external argument and draws the internal argument to the phase edge in order to check little v 's EPP feature.⁸ At the phase edge, the internal argument can check features with higher probes (namely C in this case). Just as with Tagalog v_{Tr} (Aldridge 2008), the internal argument is extracted because it is first pulled to the phase-edge by the EPP on v .

(14) Ngenó-ngené (E), Pancor (object voice)

- a. oku wah=ne_i kirim-ang surat siq loq Ali_i
 1SG PERF=3 Ø.send-APPL letter by ART Ali
 'Ali sent me a letter'



To conclude this section, despite some minor differences, the active v (N^I -) shown in (13a) is analagous to Indonesian *meN*-. Both prefixes allow verbs to project internal arguments, but these objects are unavailable for extraction due to a strong vP phase and no EPP on v to draw the object to the phase edge. For this reason, in such structures, only external arguments are able to enter into an Agree relation with higher probes. Similarly, Tagalog's antipassive v (v_{Intr}) bears no EPP to draw internal arguments to the strong phase edge, but unlike Indonesian and Sasak, Tagalog's v_{Intr} bears no case feature.

⁸ For clarity, the EPP that draws the internal argument to the phase edge is distinct from the selection of an external argument.

4. Sasak and parametricity of features on v_{Act}

4.1. Towards an account of the other Sasak nasal prefixes

Having established that the phase-based approach can account for one of the Sasak nasal prefixes, we are left with at least two questions concerning the analysis. First, does a system that allows parametric variation of single features on syntactic heads predict attested patterns, or does it overgenerate possibilities that are unattested. Secondly, is there any diachronic or synchronic evidence to support the proposal that Indonesian *meN-* may have at an earlier time been used in objectless clauses and only later acquired a case feature, thereby allowing verbs to project an additional argument? By pursuing a formal account of the other Sasak prefixes, we will arrive at an answer to the first of these questions. As for the second question, the fact that some Sasak nasal prefixes function as antipassive morphemes provides suggestive synchronic evidence to support the idea that the absence or presence of a Case feature (and therefore internal argument) is indeed one robust dimension along which the cognate nasal prefixes vary.

To begin, (8) repeated as (15) below shows how v_{Act} (realized as N^4 in Northern Sasak) must not bear a Case or EPP feature since verbs prefixed with N^4 are unable to occur with objects. As such, N^4 is analogous with Tagalog's antipassive little *v*. So, while Indonesian *meN-* and Eastern Sasak N^1 also have no EPP feature (and as a result allow extraction of only external arguments), they have acquired a Case feature and are able to occur with objects. Conversely, N^4 in Northern Sasak also has no EPP feature, but additionally no Case feature since the prefix only occurs in object-less clauses.⁹

(15) N^4 - in Northern Sasak bears neither a Case nor EPP feature

- | | | | | |
|----|-------------------------------------|-----------------|--------------|------------------|
| a. | mèlé-ng=ku | <i>menyéran</i> | kón | pawang=nó |
| | want-LNK=1SG | N^4 .hunt | place | forest=DEF |
| | ‘I want to hunt in the forest’ | | | |
| b. | *mèlé-ng=ku | <i>menyéran</i> | <i>burón</i> | kón pawang=nó |
| | want-LNK=1SG | N^4 .hunt | <i>deer</i> | place forest=DEF |
| | ‘I want to hunt deer in the forest’ | | | |

⁹ The fact that N^1 in Central/Southern Sasak and N^3 in Eastern Sasak, which also have a detransitivizing effect, may occur with semantically-restricted objects might be attributed to the fact that structural accusative Case is not always the only case that can be assigned to direct objects. Finnish accusative Case is assigned under definite readings, whereas indefinite objects are assigned (a morphologically-distinct) partitive Case (Belletti 1988). While Sasak exhibits no overt morphological case, we might instead look to cliticization patterns for evidence. N^1 in these varieties is required for non-referential patients, but an object enclitic on the verb (usually available for referential objects) is not possible with the non-referential reading (Austin 2012:20). This suggests that these non-referential patients are not licensed in the same way as the objects in clauses with bare verbs.

While N^l in Eastern Sasak appears analogous with Indonesian *meN-*, section 2 shows how northern sub-dialects vary in terms of whether or not an extraction asymmetry is observed. Gondang N^l can only occur with extracted external arguments and only the bare (non-prefixed) verb can appear with extracted internal arguments. This is just as what we find for the cognate nasal prefix in Indonesian. Conversely, data from Sukadana show that N^l appears symmetrically on verbs regardless of which argument is extracted. The examples in (6) repeated as (16) below shows that the prefix occurs on verbs over which both external (16a) and internal arguments (16b) have been extracted.

(16) N^l in Sukadana (Northern Sasak) bears both a Case and EPP feature

a. relativized actor

amaq siq *n-talet* jagung nóno pintér
 man REL N^l -plant corn that smart
 ‘The man who planted the corn is smart’

b. relativized object

jagung siq *n-talet=ku* maiq gati
 corn REL N^l -plant=1SG delicious very
 ‘The corn that I planted is very delicious’

4.2. Expanding parametric variation of features on v_{Act}

In summary, we find that the Sasak detransitivizing (or antipassive) little v (spelled-out as N^4 in Northern Sasak) presents evidence for positing a syntactic head that bears neither a Case nor EPP feature, making it equivalent to Tagalog’s antipassive v . Next, there are also data which show that N^l in some northern sub-dialects must bear a Case feature just as in Eastern Sasak and Indonesian (since an object is licensed), but in other areas it also bears an EPP feature. The presence of this EPP feature accounts for the fact that internal arguments may be extracted in the presence of the nasal prefix. After merging as a complement of the verb, these objects may raise to check v ’s EPP where they then lie at the edge of the strong vP phrase. From here they are available to check the features of higher probes, such as [uTop*] on C. These featural specifications of the syntactic heads assumed to correspond to the various Sasak nasal prefixes are given in (17) alongside the specification of Tagalog’s antipassive v and the syntactic head corresponding to Indonesian’s *meN-*.

(17) Parametric variation of the featural specification of v_{Act} in Sasak

	<i>Tagalog</i> (v_{AP})	<i>Eastern Sasak</i> (<i>& Indonesian meN-</i>)	<i>Northern Sasak</i>	
		N^l	N^l	N^4
[uCase: ACC]	no	yes	yes	no
[uD*] feature	no	no	yes	no

While the Sasak data can be accounted for by assuming that parametric variation can target individual features on syntactic heads (namely little v in this case, as first proposed by Aldridge (2008)), there is one logically possible combination not present in (17). This is a scenario in which v_{Act} bears an EPP, but no Case feature. However, the absence of an internal argument may simply render the question of an EPP moot. In other words, if a verb cannot project an internal argument, what would it mean to say that the verb does not exhibit an extraction asymmetry? Whether or not this is of serious detriment to the analysis, and hence the type of parametric variation considered in this paper, is not an issue that will be resolved in this paper. Nonetheless it represents an issue that deserves further attention.

5. Conclusion and further questions

Dialectal variation in the uses of the Sasak nasal prefixes has proven to be prime testing ground for both what connections hold between verbal morphology and syntax, as well as Aldridge's (2008) phase-based proposal that language variation may target single features present on a syntactic head. This paper has shown how the syntactic variation found with the various Sasak nasal prefixes can be described primarily along two dimensions: (1) whether the nasal verb projects an internal argument, and (2) whether or not an internal argument can be extracted.

While previously untested, allowing parametricity to target single features on syntactic heads (at least in the case of little v) allows us to account for the multiple Sasak nasal prefixes. Unlike Indonesian *meN-* (and some Sasak nasal prefixes), which must bear structural case in order to license an internal argument, some Sasak nasal prefixes function as antipassive markers. As such, they must bear no structural case, and provide suggestive synchronic evidence in favor of the idea that the Indonesian cognate nasal prefix may have originated from an antipassive marker. Furthermore, the presence of absence of an EPP feature allows an account of the (a)symmetries observed in different dialects.

There are at least two remaining questions concerning the analysis offered in this paper. First, the end of section 4 discussed one logically possible, but unattested, featural specification of little v . Secondly, while the properties of the prefixes have been discussed as static, there may be reason to question the stability of the syntactic properties of the prefixes under discussion in this paper. This is particularly true given the variability reported in Cole, Hermon and Yanti's (2008) discussion of Malay dialects, as well as optionality reported by one Northern Sasak speaker noted in section 2.

References

- Aldridge, Edith. 2008. Phase-based account of extraction in Indonesian. *Lingua* 118:1440-1469.
- Aldridge, Edith. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122:192-203.
- Asikin-Garmager, Eli. 2015. *Kutó-kuté Sasak morphosyntax (Sukadana, North Lombok)*. Unpublished manuscript.
- Austin, Peter K. 1998. *Working papers in Sasak, Vol 1*. Melbourne: University of Melbourne.
- Austin, Peter K. 2000. *Working papers in Sasak, Vol 2*. Melbourne: University of Melbourne.
- Austin, Peter K. 2012. Too many nasal verbs: Dialect variation in the voice system of Sasak. *Voice variation in Austronesian languages of Indonesia*, ed. A. Adelaar, 29-46. *NUSA* 54.
- Belletti, Adriana. 1988. The case of unaccusatives. *Linguistic Inquiry* 19(1):1-34.
- Chomsky, Noam. 2001. Derivation by phase. *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Cole, Peter, Gabriella Hermon, and Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118:1500-1553.
- Cooreman, Ann. 1994. A functional typology of antipassives. *Typological studies in language, Vol 27, Voice: form and function*, ed. Barbara A. Fox and Paul J. Hopper, 49-88. Philadelphia, PA: John Benjamins Publishing Company.
- Rickowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36(4):565-599.
- Shibatani, Masayoshi. 2008. Relativization in Sasak and Sumbawa, Eastern Indonesia. *Language and Linguistics* 9.4:865-916.
- Shibatani, Masayoshi, and Ketut Artawa. 2003. The middle voice in Balinese. *SEALS XIII: Papers from the 13th Meeting of the Southeast Asian Linguistics Society*, ed. Iwasaki, Shoichi, Andrew Simpson, Karen Adams, and Paul Sidwell, 239-261. Canberra: Pacific Linguistics.
- Wouk, Fay. 1999. Sasak is different: A discourse perspective on voice. *Oceanic Linguistics* 38(1):91-114.

SLUICING IN TAGALOG: STRATEGIES AND IMPLICATIONS*

Lena Borise
Harvard University
borise@fas.harvard.edu

This paper reconsiders the approach to Tagalog sluicing developed in Kaufman & Paul (2006) and Kaufman (2006), and puts forward an alternative analysis. I propose that Tagalog has two distinct strategies for sluicing that follow the two wh-question formation strategies available in the language: pseudoclefts for argument wh-questions, and wh-movement for adjunct wh-questions. Such a bifurcation is problematic for the traditional approaches to sluicing. I therefore propose that the Tagalog data discussed here provides support for the Unconstrained Pseudosluicing Hypothesis as argued for in Barros (2014).

1. Introduction

The aim of this paper is to investigate sluicing—a type of TP-ellipsis introduced by a wh-word—in Tagalog. It is generally accepted that in English and other languages that have wh-movement, sluicing relies on the embedded wh-question formation strategy (Ross, 1969). Tagalog, however, has two distinct wh-question formation strategies. Therefore, better understanding of the nature of sluicing in a language like Tagalog is essential for the theory of sluicing.

The structure of this paper is as follows. Section 2 summarises previous research on sluicing and sluicing-like phenomena. Section 3 shows that the approach to Tagalog sluicing put forward in Kaufman & Paul (2006) and Kaufman (2006) is not supported by data. Section 4 introduces an alternative analysis. Section 5 concludes.

2. Sluicing: Theoretical Background

The notion ‘sluicing’ goes back to the seminal paper by Ross (1969), and has since been used as a cover term for the type of TP-ellipsis with a wh-remnant in an embedded clause, as in (1):

- (1) *Somebody left the door open, but I don't know who.*

* I would like to thank Maria Polinsky, Brooke Larson, Adam Szczegielniak, and Daniel Kaufman for their advice on this project, as well as the audiences at Polinsky Lab, BLS 41 and AFLA 22 for their most helpful feedback. Special thanks are due to the Tagalog speakers that provided the data used in this paper: Beverly Ho, Henrison Hsieh, Jennifer Tan, and two other informants who chose to remain anonymous.

The original analysis proposed by Ross (1969) and widely adopted afterwards states that in English sluicing formation parallels embedded wh-question formation. Namely, in both, the wh-word in the embedded clause is raised to Spec-CP; in sluicing, it is followed by ellipsis of the embedded TP:

- (2) [TP *Somebody* [T' *sneezed*], but I don't know [CP *who*_i [TP *t*_i *sneezed*]].

There are certain structural parts that all sluicing examples have in common: the correlate in the antecedent clause is matched by the wh-word remnant in the sluice, and the combination of remnant and ellipsis site, under Ross (1969) approach, is structurally parallel to the antecedent clause:

- | | | | | | |
|-----|-------------------|----------------|---------------------------|------------|---------------------|
| (3) | [<i>Someone</i> | <i>sneezed</i> | [<i>but I don't know</i> | <i>who</i> | [<i>sneezed</i>]] |
| | correlate | | | remnant | ellipsis site |
| | antecedent clause | | | sluice | |

Sluicing is often contrasted with pseudosluicing, the notion introduced in Merchant (1998) to describe the sluicing-like construction such as the one in Japanese shown below:

- (4) Dareka-ga sono hon-o yon-da ga, [watashi-wa [pro dare
 someone-NOM that book-ACC read-PST but 1sg-TOP pro who
 datta ka] wakaranai].¹
 be.PST Q know.NEG
 Someone read that book, but I don't know who.

(from Merchant, 1998)

The main difference between the two phenomena lies in the fact that in pseudosluicing the ellipsis site contains a copular clause. In a copular clause, the predicate is a non-verbal category, such as a DP, and there may or may not be an overt copula introducing the predicate. In pseudosluicing, therefore, the ellipsis site is not syntactically parallel to the antecedent clause.

Instances of sluicing in which there is no overt correlate in the antecedent clause have been dubbed 'sprouting' (Chung, Ladusaw, and McCloskey, 1995). In English, sprouting is possible both with argument and adjunct sluices:

- (5) Mary [T' is eating], but I don't know [CP what [TP *t*_i]].
- (6) John [T' plays tennis on Sundays], but he didn't say [CP where [TP *t*_i]].

¹ Abbreviations used in the glosses: 1pl—1st person plural; 1sg—1st person singular; 2sg—2nd person singular; 3sg—3rd person singular; ACC—accusative; AV—actor voice; CAUS—causative; COMP—complementiser; DAT—dative; DIR—directive; EXCL—exclusive; EXT—existential; GEN—genitive; INT—intensive; IPFV—imperfective; IRR—irrealis; LNK—linker; LV—locative voice; NEG—negation; NOM—nominative; PFV—perfective; PST—past; PV—patient voice; RED—reduplication; Q—question particle; TOP—topic.

As you can see, *what* in (5) and *where* in (6) have no overt correlate in the antecedent clause. A salient property of sprouting is its island sensitivity - sprouting within an island leads to ungrammaticality:

- (7) *Sandy is very anxious to see which students will be able to solve the homework problem, but she won't say how.

(from Chung, Ladusaw, and McCloskey, 1995)

In the literature on sluicing, discussion revolves around three large issues: (i) sluicing-licensing conditions, (ii) interaction of sluicing with island constraints, and (iii) distinction between sluicing and pseudosluicing. In the remainder of this section, I will briefly introduce each of these questions.

As a type of ellipsis, sluicing should be licensed by some sort of identity between the antecedent clause and the ellipsis site (cf. Lasnik, 2001; Merchant, 2001, 2005). There is a debate in the literature as to whether sluicing is licensed by syntactic or semantic identity between the sluice and the antecedent clause. Under the syntactic view, advocated by Ross (1969), Chung, Ladusaw, and McCloskey (1995), Merchant (2008), and others, sluicing is licensed if the antecedent clause and the sluice have matching syntactic structure. Under the semantic approach, put forward in Baker and Brame (1972), Merchant (2001), Abels (2011), and others, it is mutual entailment between the antecedent clause and the sluice that licenses sluicing. There are also hybrid syntax-semantics proposals, like Barros (2014) and Chung (2014). Since the Tagalog data discussed in this paper suggests that there are two sluicing strategies in Tagalog, it also suggests that neither exclusively syntactic nor exclusively semantic sluicing-licensing condition is sufficient on its own, in line with the hybrid proposals.

Various claims have been made about the nature of interaction between sluicing and island constraints. Ross (1969) points out that sluicing has the property of ameliorating islands, based on examples like (8) and (9) below. While (8), a sluicing construction, is judged as acceptable, the non-elided version involving wh-movement from an island in (9) is not:

- (8) She kissed a man who bit one of my friends, but Tom doesn't realise
[_{CP} which one of my friends [_{TP} t_i]].
- (9) *She kissed a man who bit one of my friends, but Tom doesn't realise
[_{CP} which one of my friends]_i [_{TP} she kissed a man who bit t_i]].

In more recent work the hypothesis that sluicing ameliorates islands has been challenged. Specifically, it has been proposed (e.g., Abels, 2011; Barros, 2014) that sluicing evades islands instead of ameliorating them, by resorting to a non-isomorphic structure in the ellipsis site. The only exception to this is presented by cases of contrast sluicing (Abels, 2011), which will not be discussed here.

Finally, the relationship between sluicing and pseudosluicing has been a matter of debate. In Merchant's (1998) original proposal for Japanese the defining

difference between sluicing and pseudosluicing is that pseudosluicing is taken to be an instance of copula drop combined with pronoun drop. Both of these phenomena are independently available in languages like Japanese, and a combination of the two on the surface produces a result similar to sluicing. However, since pseudosluicing, unlike sluicing, is not derived by TP-ellipsis, it has properties different from sluicing - for instance, Japanese pseudosluicing is sensitive to islands. Based on this and other facts, Merchant (1998) argues that pseudosluicing does not exist in English.

However, the notion of pseudosluicing has since evolved to apply to a broader set of constructions, including cases when the ellipsis site in English sluicing is analysed as a copular clause, and not as ellipsis of a TP that is syntactically isomorphic to the antecedent clause (e.g., Rodrigues et al., 2009). Subsequently it has been argued that a non-isomorphic structure of the sluice is employed in cases where the isomorphic one incurs an island violation - as already discussed above in the context of interaction between sluicing and island constraints. Under this view, the underlying structure of (8) is not (9) but (10) below:

- (10) She kissed a man who bit one of my friends, but Tom doesn't realise
[_{CP} which one of my friends]_i [_{TP} that was]].

Under this approach, pseudosluicing exists in English but has a rather restricted distribution - it is only employed in cases in which sluicing would yield an ungrammatical structure. Such an approach was dubbed Constrained Pseudosluicing Hypothesis by Barros (2014).

It is in contrast with Barros' own Unconstrained Pseudosluicing Hypothesis. As the name suggests, under this latter hypothesis, not only does pseudosluicing exist in languages like English, but it also has much broader distribution. Specifically, Barros (2014) argues that in cases where the ellipsis site in sluicing might legitimately contain a sluicing as well as a pseudosluicing structure, both are possible. According to the Unconstrained Pseudosluicing Hypothesis, both (12) and (13) are examples of possible underlying structures for (11):

- (11) Somebody [_{T'} left the door open], but I don't know who.
(12) Somebody [_{T'} left the door open], but I don't know [_{CP} who_i [_{TP} t_i left the door open]].
(13) Somebody [_{T'} left the door open], but I don't know [_{CP} who_i [_{TP} t_i that was]].

The Tagalog data presented in this paper lends additional support to the Unconstrained Pseudosluicing Hypothesis, due to the facts that: (i) Tagalog can have pseudosluicing even when there is no island in the antecedent clause, and (ii) there is substantial structural similarity between sluicing and pseudosluicing in Tagalog.

Note that in this paper, the term pseudosluicing is used in the Unconstrained Pseudosluicing Hypothesis sense rather than in Merchant's original sense: it is taken to be a type of TP-ellipsis in a copular clause, not an instance of copula drop combined with pronoun drop.

More recently, sluicing-like phenomena in languages other than English have received increasing attention. Since it is generally agreed that sluicing in English relies on the embedded *wh*-question strategy, it raises the question of whether languages that do not have *wh*-movement of the English type have sluicing. Sluicing-like structures exist in most languages, and for many it is argued that the sluicing strategy relies on the *wh*-question formation strategy. Specifically, it has been proposed for Persian (Toosarvandani, 2008) and Georgian (Erschler, 2015), in which *wh*-movement targets the focus projection below CP, that sluicing is derived in the same way. For languages with *wh*-in-situ, like Japanese (Merchant, 1998) and Mandarin Chinese (Adams and Tomioka, 2012), a pseudosluicing analysis has been proposed. For languages that employ pseudoclefts to form *wh*-questions, like Malagasy (Potsdam, 2007), it is argued that the pseudocleft strategy is also used to form sluicing. With this in mind, let us turn to Tagalog data.

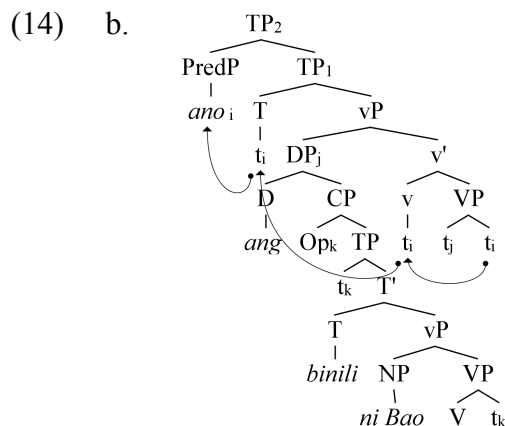
3. Tagalog

3.1. Wh-question Formation in Tagalog

As already mentioned, it is generally agreed that sluicing utilises the language-specific *wh*-question formation strategy. Tagalog has two distinct *wh*-question formation strategies: pseudoclefts for argument *wh*-questions, and *wh*-movement of the English type for adjunct *wh*-questions (Aldridge, 2002). Argument *wh*-questions therefore consist of a clause-initial *wh*-word that acts as a predicate nominal, and a headless relative clause following it; as with other non-verbal predicates, there is no overt copula:

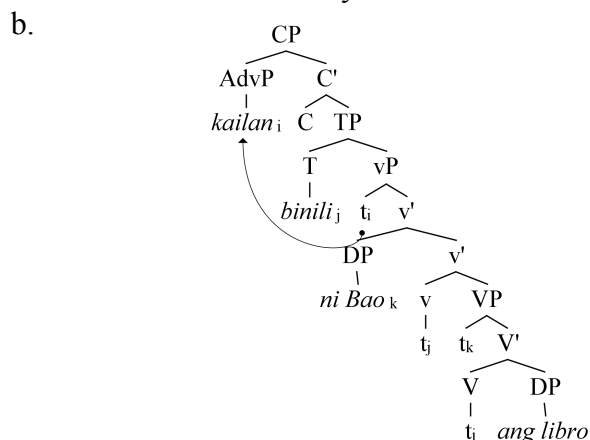
- (14) a. [_{TP} Ano [_{DP} *(ang) [_{CP} b<in>ili ni Bao?]]]
 what NOM <PV.PFV>buy GEN Bao
 What did Bao buy?

In (14b), I take the topmost TP₂ to be equivalent to the functional projection FP that Potsdam (2007) postulates for Malagasy sluicing. In contrast with Malagasy, however, Tagalog subjects stay in the *v*P; based on this, a simpler alternative to (14b) would have a single TP, TP₁, with *ano* sitting in its specifier. In that case, however, we would have to say that sluicing elides the *v*P. This contradicts what we know about ellipsis in general (the parts of the tree that can undergo ellipsis are TP, VP and NP), and about sluicing in particular (sluicing elides the TP). Therefore, I take there to be two TPs in (14b):



In contrast with argument wh-questions, adjunct wh-questions are formed by fronting to Spec-CP of the wh-word that does not act as a nominal predicate:

- (15) a. [_{CP} Kailan (*ang) b<in>ili ni Bao ang libro?]
 when NOM <PV>buy GEN Bao NOM book?
 When did Bao buy the book?



One of the main pieces of evidence supporting this distinction comes from the distribution of the case marker *ang*, which is used to mark the subject in Tagalog. In the argument wh-question in (14) *ang* cannot be omitted, which signals that the headless relative CP that *ang* takes as a complement acts as the subject of the clause, whereas *ano* ‘what’ acts as the predicate. In contrast with (14), in the adjunct wh-question in (15) *ang* cannot appear.

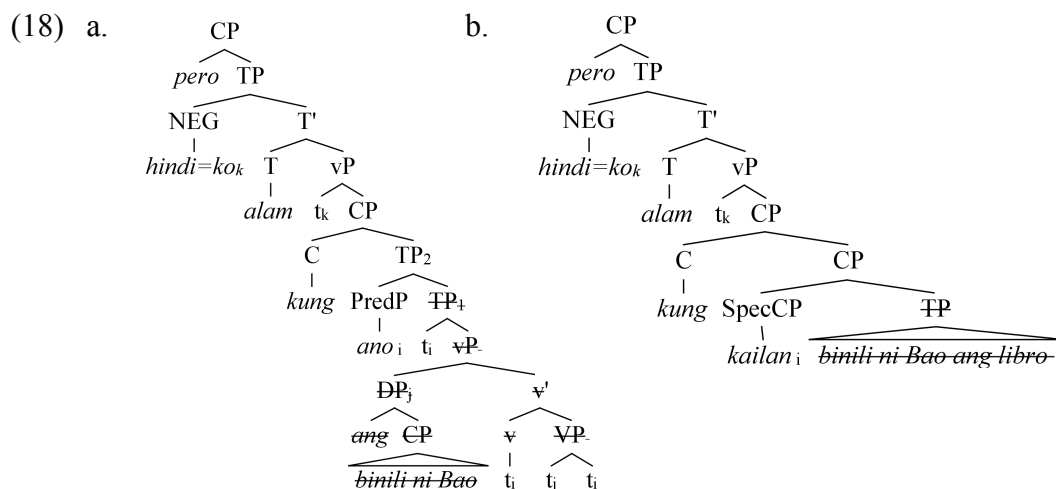
These facts suggest that argument and adjunct wh-questions in Tagalog have different structures: argument wh-questions are biclausal, with the wh-word acting as a nominal predicate and taking a headless relative clause as a subject; adjunct wh-questions, on the other hand, are monoclausal structures involving wh-movement.

The distinction between the two structures is also confirmed by clitic placement. Aldridge (2002) notes that certain personal pronoun clitics (as well as

some aspectual clitics - Richards, p.c.) can only attach to wh-words that undergo wh-movement. Since argument wh-words undergo head-raising, clitics are predicted to be housed only by adjunct wh-words. This is confirmed both for wh-questions (see Aldridge, 2002) and sluicing:

- (16) [CP P<um>unta =ka sa Maynila], [CP pero naka-limut-an
 <PFV.AV>go 2sg.NOM DAT Manila but PFV.BV-forget-DIR
 ko [CP kung kalian =ka p<um>unta]]²
 1sg COMP when 2sg.NOM <PFV.AV>go
 You went to Manila, but I forgot when (you went).
- (17) [CP May <g<in>a>gawa =ka, [CP pero hindi ko alam
 EXT <RED<PV>>do 2sg.NOM but NEG 1sg know
 [CP kung ano (*=mo) [TP ang (*=mo) <g<in>a>gawa *(=mo)]]]]³
 COMP what 2sg.GEN NOM 2sg.GEN <RED<PV>>do 2sg.GEN
 You did something, but I don't know what you did.

Based on these distributional facts, and also taking into account the crosslinguistic evidence, I propose that Tagalog has two sluicing strategies too, corresponding to the argument-adjunct distinction found in wh-questions. The corresponding sluicing structures then look the following way: (18a) for arguments, (18b) for adjuncts.



3.2. Sluicing in Tagalog

Sluicing in Tagalog has not received much attention in the literature so far. Some ideas about the possible analysis can be found in Kaufman & Paul (2006). Kaufman

² The clitic cannot be final in a clause, so in cases of sluicing with clitics, the verb following the clitic needs to be spelled-out.

³ *ka/mo* alternation is due to the voice change in the existential when embedded.

and Paul observe that sluicing in Tagalog has two cross-linguistically unusual properties: (i) ban on sprouting in argument sluices, and (ii) sensitivity to certain types of islands. However, neither of these turn out to be supported by data.

3.2.1. Sprouting with argument sluices

Kaufman & Paul (2006) provide examples showing that sprouting is banned in argument sluices:

- (19) *_{[CP} K<um>akanta si Maria] _{[CP} pero hindi ko alam kung ano].
<AV>sing.IPFV NOM Maria but NEG 1sg know COMP what
Intended: Maria was singing but I don't know what.

In adjunct sluices, however, sprouting is allowed:

- (20) _{[CP} K<um>akanta si Maria] _{[CP} pero hindi ko alam kung saan].
<AV>sing.IPFV NOM Maria but NEG 1sg know COMP where
Maria was singing but I don't know where.

(from Kaufman & Paul, 2006)

Kaufman (2006) accounts for the lack of sprouting in argument sluicing by proposing that Tagalog verbs are ambitransitive: that is, unless the internal argument of the verb is introduced overtly, the verb is intransitive. Under this hypothesis, (19) is degraded because of the transitivity mismatch between the matrix verb and the verb within the elided part of the embedded clause, much like it is with certain English verbs:

- (21) a. She bathed someone, but I'm not sure whom.
b. *She bathed, but I'm not sure whom.

Kaufman's (2006) approach also accounts for the fact that sprouting is allowed with adjuncts, since adjuncts do not interact with the transitivity of the verb.

However, judgments vary in cases like (19), which was noticed both by Kaufman (p.c.) and myself. Specifically, there is a division among speakers when it comes to sluices with indefinite argument correlates, as in (19).

While marked as unacceptable in Kaufman & Paul (2006), and as marginal in Kaufman (2006), (19) is in fact accepted and even preferred by some Tagalog speakers. Specifically, out of the four informants I checked (19) and similar examples with, two allow it, and one prefers it to the alternative (the fourth informant does not accept (19) in either Tagalog or English). The alternative to (19) introduces the indefinite argument correlate by means of the existential construction with *may*, an impersonal predicate used to introduce indefinite arguments (Keenan, 2009; Aldridge, 2012):

- (22) [CP May k<in>anta si Kim sa handaan] [CP pero hindi ko alam
EXT <PV>sing NOM Kim DAT party but NEG 1sg know
[CP kung ano.]]
COMP what
Kim sang something at the party, but I don't know what.

Moreover, for the two speakers that allow both (19) and (22), the two variants are distinct in terms of their information structure. For these speakers, (19) bears broad focus - that is, it is a suitable reply to a question 'What happened?', while (22) is a structure with narrow focus on the object - i.e., it is a suitable reply to a question 'What did Kim sing?'

The hypothesis that the existential construction brings the correlate in focus is supported by the speakers' intuition that (23) feels 'disjointed', similarly to its English counterpart:

- (23) # [CP May in-inom si Kim sa handaan] [CP pero hindi ko alam
EXT PV-drink NOM Kim DAT party but NEG 1sg know
kung kailan.]
COMP when
#There was something that Kim drank at the party, but I don't know when.

The focal properties of both Tagalog constructions require further investigation. The overall conclusion, however, is that, at least for some Tagalog speakers, sprouting with arguments does not pose any problems.

3.2.2. Island sensitivity

To illustrate that Tagalog sluicing is sensitive to islands, Kaufman & Paul (2006) use an example of a relative clause island. They claim that both argument and adjunct sluices are sensitive to island constraints:

- (24) *[CP Kilala ni Kim ang tao [TP -ng may b<in>abasa]]
know GEN Kim NOM person-LNK EXT <PV>read.IPFV
[CP pero hindi niya sasabih-in [CP kung ano.]]
but NEG 3sg say.IRR-PV COMP what
Intended: Kim knows the person who was reading (something) but she won't say what.
- (25) *[[CP Kilala ni Kim ang tao [TP -ng may b<in>abasa]]
know GEN Kim NOM person-LNK EXT <PV>read.IPFV
[CP pero hindi niya sasabih-in [CP kung saan.]]]
but NEG 3sg say.IRR<PV> COMP where
Intended: Kim knows the person who was reading (something) but she won't say where.

(from Kaufman & Paul, 2006)

Unlike relative clauses, coordinate structure islands, adjunct islands, and complex NP islands do not block sluicing - as shown in (26), (27), and (28), respectively. This data brings Tagalog sluicing in line with cross-linguistic evidence.

- (26) [_{CP} Nag-luto si Ben ng menudo at ng isa pa-ng
<PFV.AV>cook NOM Ben GEN menudo and GEN one still-LNK
putahe] [_{CP} pero hindi ko alam [_{CP} kung alin.]]
dish but NEG 1sg know COMP which
Ben cooked menudo and another dish, but I don't know which.

(from Kaufman & Paul, 2006)

- (27) [_{CP} Um-uwi si Kim [_{CP} dahil may kailangan siya-ng gaw-in]],
AV-go.home NOM Kim because EXT must 3sg-LNK do-PV
[_{CP} pero hindi ko alam [_{CP} kung ano.]]
but NEG 1sg know COMP what
Kim went home because she needed to do something, but I don't know what.

- (28) [_{CP} Na-rinig ni Kim ang balita [_{CP} na in-ayos ni Fred
PFV.PV-hear GEN Kim NOM news LNK PV-sort.PFV GEN Fred
ang problema]], [_{CP} pero hindi niya alam [_{CP} kung alin]].
NOM problem but NEG 3sg know COMP which
Kim heard the news that Fred solved a problem, but she doesn't know which.

At first glance, these facts suggest that the only type of islands that interacts with sluicing in Tagalog is a relative clause island, as shown in (24) and (25) above. However, it can easily be shown that a relative clause on its own is not an obstacle for sluicing:

- (29) [_{CP} Alam ni Bao ang babae [_{TP} -ng nag-susulat ng libro],
know GEN Bao NOM woman-LNK INT-write.AV GEN book
[_{CP} pero hindi niya maalala [_{CP} kung bakit.]]
but NEG 3sg remember.PV COMP why
Bao knows a woman that was writing a book, but he doesn't remember why.⁴

Another plausible hypothesis is that it is the existential construction embedded in the relative clause that interferes with sluicing, but, surprisingly, other examples of the same structure (relative clause embedding the existential) do not

⁴ It has been noted (Szczegielniak, p.c.) that (29) is an example of sprouting within an island, and as such, should be ungrammatical, as shown in (7). Manipulating the lexical items - specifically, introducing a different subject in the 'but...' clause indeed degrades the acceptability both in English and Tagalog - cf. *Bao knows a woman that was writing a book, but Fred doesn't remember why*. For now, I am leaving this question open.

block sluicing, as shown in (30). It should be noted that, due to the overall complexity of (30), caused by multiple embeddings, the speakers find it easier to process when the ellipsis site is spelled out.

- (30) [CP Naka-salubong ko ang isa-ng babae [CP na [TP may
PRF.AV-run_into 1sg NOM one-LNK girl LNK EXT
h<in>ahanap sa kanya-ng pitaka]]] [CP pero hindi ko alam
<IPRF.PV>look_for DAT her-LNK purse but NEG 1sg know
[CP kung ano (ang h<in>ahanap niya)]]
COMP what NOM <IPRF.PV>look_for 3sg
I ran into a girl that was looking for something in her purse, but I don't
know what (she was looking for).

Overall, as we have seen, there is no compelling evidence that Tagalog sluicing is sensitive to island constraints. As for the unacceptable examples (24-25), which none of my consultants accepted either, it seems that the reason for degraded judgments has to do with the difficulty of establishing pronominal reference, and not sluicing as such.

4. An Alternative Analysis

4.1. Sluicing and Pseudosluicing

Another observation due to Kaufman & Paul (2006) is that an overt demonstrative pronoun, *iyon* 'that' can surface as the subject of the embedded clause in argument sluices. Such cases constitute pseudosluicing structures with the ellipsis site spelled out in full (31). *Iyon*-insertion is impossible in adjunct sluices (32), similarly to the impossibility of the English counterpart to be followed by *that was*:

- (31) [CP May binigy-an ng pera], [CP pero hindi namin alam
EXT give.PFV-LV GEN money but NEG 1pl.EXCL.GEN know
[CP kung sino [TP iyon.]]]
COMP who that

Someone was given money but we don't know who it was.

- (32) [CP Na-i-pa-ayos niya ang kotse] [CP pero hindi ko alam
PFV-PV-CAUS-fix 3sg.GEN NOM car but NEG 1sg know
[CP kung paano (*iyon).]]
COMP how (that)

She fixed the car but I don't know how (*that was).

(from Kaufman & Paul, 2006)

Recall that argument sluices have a pseudocleft structure, as shown in (18a). Pseudosluicing with arguments, as in (31), is a copular clause. In the remainder of this section I show that in Tagalog pseudoclefting and pseudosluicing are minimally different implementations of the same structure.

employ the pseudoclefting strategy, which is also what is used in copular clauses in pseudosluicing.

Adopting such an analysis and differentiating argument and adjunct sluices also allows to account for the fact that *iyon* can only appear in argument sluices: since adjunct wh-words do not act as nominal predicates, they cannot take *iyon* as the subject.

Furthermore, if (31) and (33) are interchangeable subtypes of the same predicate-subject structure, they lend support to the Unconstrained Pseudosluicing Hypothesis, as opposed to the Constrained Pseudosluicing Hypothesis. Recall that according to the latter, pseudosluicing can only be employed in cases where regular sluicing would run into an island constraint. The Tagalog data, however, shows that both sluicing and pseudosluicing employ a pseudocleft structure. Therefore, they can be used interchangeably, even in cases where no island is involved. This is predicted to be impossible by the Constrained Pseudosluicing Hypothesis, and therefore supports the Unconstrained Pseudosluicing Hypothesis.

Moreover, the very fact that the two structures in (31) and (33) are so similar structurally (much more so than the respective English sluicing and pseudosluicing counterparts) is consistent with the intuition that they should be treated as two instances of the same construction.

4.2. The Existential Construction and Sprouting

Recall that sprouting within an island has been argued to lead to ungrammaticality (Chung, Ladusaw, and McCloskey, 1995), as shown for English in (7), repeated below:

- (7) *Sandy is very anxious to see which students will be able to solve the homework problem, but she won't say how.

It is therefore surprising that examples like (30), as repeated below, are judged as acceptable. Since there is no overt correlate for the wh-word *ano* in the matrix clause, they also constitute examples of sprouting within an island.

- (30) [CP Naka-salubong ko ang isa-ng babae [CP na [TP may
PRF.AV-run_into 1sg NOM one-LNK girl LNK EXT
h<in>ahanap sa kanya-ng pitaka]]] [CP pero hindi ko alam
<IPRF.PV>look for DAT her-LNK purse but NEG 1sg know
[CP kung ano (ang h<in>ahanap niya)]]
COMP what NOM <IPRF.PV>look_for 3sg
I ran into a girl that was looking for something in her purse, but I don't
know what (she was looking for).

The reason for the fact that there is no overt correlate in (30) is that in Tagalog, indefinite arguments like *something* or *someone* must be introduced by means of the existential *may*. Tagalog speakers have a strong intuition that *may* is a

counterpart of *something*, like in cases with argument sprouting, discussed in Section 3.2.1. Compare (22) with (19):

- (22) [CP *May* k<in>anta si Kim sa handaan] [CP pero hindi ko alam
EXT <PV>sing NOM Kim DAT party but NEG 1sg know
[CP kung ano.]]
COMP what
Kim sang *something* at the party, but I don't know what.
- (19) *[CP K<um>akanta si Maria] [CP pero hindi ko alam kung ano].
<AV>sing.IPFV NOM Maria but NEG 1sg know COMP what
*Maria was singing but I don't know what.

It seems therefore that in examples like (30) and (22) there is no sprouting - instead, the existential verb *may* plays the role of the non-nominal correlate for the wh-remnant. To the best of my knowledge, there are no instances of non-nominal correlates discussed in the literature. The Tagalog case, therefore, requires further investigation.

5. Conclusion

To conclude, this paper argued that out of the three possible analyses for Tagalog sluicing that have been previously considered in the literature, the most promising one is that there are two sluicing strategies in Tagalog: wh-movement for adjunct sluices, and a pseudocleft structure for argument sluices. In this way, Tagalog evidence is in line with the cross-linguistic generalisation that sluicing formation utilises the language-specific wh-question formation strategy.

Also, I have shown that Tagalog is cross-linguistically unusual in that the sluicing and pseudosluicing constructions in the language are instances of the same underlying pseudocleft structure. The difference between sluicing and pseudosluicing lies only in the complexity of the subject DP. These facts also support the Unconstrained Pseudosluicing Hypothesis.

Like sluicing in other languages, Tagalog sluicing is not sensitive to island constraints. As for the sprouting asymmetry that has been reported before—sprouting being possible with adjunct but not argument sluices—it has not been confirmed by the data, though there is some inter-speaker variation. The hypothesis put forward in this paper is that sprouting in argument sluices differs from the alternative construction headed by the existential *may* in terms of its discourse properties: the former has broad focus, and the latter bears narrow focus.

Since this paper proposes that there are two sluicing strategies in Tagalog, it also suggests that neither exclusively syntactic nor exclusively semantic sluicing-licensing condition is sufficient on its own. Because argument sluices have pseudocleft structure, there is no syntactic parallelism between the sluice and the antecedent clause. This is problematic for the syntactic condition. As for the semantic

condition, it does not allow to incorporate the generalisation that argument and adjunct sluices in Tagalog have different structures, because the semantic condition does not take syntax into account. Therefore, the Tagalog sluicing data calls for a hybrid syntax-semantic condition. Determining its nature is a task for future work.

References:

- Abels, Klaus. 2011. Don't repair that island! It ain't broke. Paper presented at Islands in Contemporary Linguistic Theory. University of the Basque Country, Spain.
- Adams, Perng Wang, and Satoshi Tomioka. 2012. Sluicing in Mandarin Chinese: An instance of pseudo-sluicing. *Sluicing: Cross-Linguistic Perspectives*, eds. Andrew Simpson and Jason Merchant, 219-247. Oxford University Press
- Aldridge, Edith. 2002. Nominalization and WH-movement in Seediq and Tagalog. *Language and Linguistics* 3.2:393-426.
- Aldridge, Edith. 2012. Event Existentials in Tagalog. *Proceedings of AFLA 18*, 16-30. Asia-Pacific Linguistics.
- Baker, Carl L., and Michael K. Brame. 1972. 'Global rules': a rejoinder. *Language* 48.1:51-75.
- Barros, Matthew. 2014. *Sluicing and Identity Ellipsis*. Doctoral Dissertation, Rutgers University.
- Chung, Sandra, William A. Ladusaw, and James McCloskey. 1995. Sluicing and logical form. *Natural Language Semantics* 3.3:239-282.
- Chung, Sandra. 2014. Syntactic Identity in Sluicing: How Much and Why. *Linguistic Inquiry* 44.1:1-44.
- Erschler, David. 2015. *Embedded Questions and Sluicing in Georgian and Svan*. lingbuzz/002410, accessed on 18th April 2015.
- Kaufman, Daniel. 2006. Transitivity in Austronesian and sluicing. Paper presented at International Conference on Austronesian Linguistics (ICAL) X. Puerto Princesa, Philippines.
- Kaufman, Daniel & Paul, Ileana. 2006. Sluicing in Tagalog. Paper presented at the Austronesian Formal Linguistics Association (AFLA) XIII. National Tsing Hua University, Taiwan.
- Lasnik, Howard. 2001. When can you save a structure by destroying it? *Proceedings of NELS* 31 (2), 301-320. Georgetown University.
- Merchant, Jason. 1998. 'Pseudosluicing': elliptical clefts in Japanese and English. *ZAS Working Papers in Linguistics* 10: 88-112.
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford University Press.
- Merchant, Jason. 2005. Fragments and ellipsis. *Linguistics and philosophy* 27.6: 661-738.
- Merchant, Jason. 2008. An asymmetry in voice mismatches in VP-ellipsis and pseudogapping. *Linguistic Inquiry* 39.1: 169-179.

- Potsdam, Eric. 2007. Malagasy sluicing and its consequences for the identity requirement on ellipsis. *Natural Language & Linguistic Theory* 25(3), 577-613.
- Rodrigues, Cilene, Andrew Nevins, and Luis Vicente. 2009. Cleaving the interactions between sluicing and preposition stranding. *Romance languages and linguistic theory*, eds. L. Wetzels and J. van der Weijer, 175–198. John Benjamins.
- Ross, John Robert. 1969. Guess who? *Proceedings of CLS 5*. University of Chicago.
- Toosarvandani, Mazyar. 2008. *Wh*-movement and the syntax of sluicing. *Journal of Linguistics* 44: 677-722.

THREE WAYS TO “STEAL” AN ELEMENT FROM A CP: EVIDENCE FROM FORMOSAN

Victoria Chen
University of Hawai‘i
ychen6@hawaii.edu

Shin Fukuda
University of Hawai‘i
fukudash@hawaii.edu

Raising-to-object (RTO) refers to a phenomenon in which a semantically and thematically embedded element (an XP) appears to surface outside of the embedded clause and exhibits structural characteristics of matrix objects. Past studies have shown that RTO constructions cross-linguistically can be divided into at least six subtypes, which vary in terms of the size of the embedded complement and the presence or absence of a movement of an XP out of the embedded clause. This paper presents novel RTO data from three Formosan languages, Puyuma, Amis, and Seediq, which exhibit a number of similarities, yet differ from one another in terms of (i) the case restriction of the XP and (ii) island-sensitivity. We argue that these micro-variations can be accounted for by analyzing RTO in the three languages as utilizing three distinct strategies that diverge from one another in terms of the presence or absence of a null Operator (Op) inside the embedded clause, and the status of the Op as being base-generated or A'-moved. If the present analysis is on the right track, the micro-variation in RTO in these three Formosan languages provides compelling evidence for the recent proposal that CPs are not natural predicates and that languages may utilize different strategies in establishing a relationship between a CP and a constituent external to it (Rothstein 1991; Landau 2011).

1. Introduction

Raising to Object (RTO) refers to a phenomenon in which a semantically and thematically embedded element appears to surface outside of the embedded clause and exhibits characteristics typical of matrix objects (e.g., case marking, agreement, and binding). This paper analyzes apparent cases of RTO in Puyuma (1a), Amis (1b), and Seediq (1c), three Formosan languages that belong to three primary branches of the Austronesian family and exhibit highly productive RTO constructions.

* This project is supported by funding from the Department of Linguistics at University of Hawai‘i at Mānoa and the Linguistics Institute of Academia Sinica. We would like to thank Robert Blust, Kamil Deen, Ying-chin Lin, Diane Massam, William O’Grady, Yuko Otsuka, Masha Polinsky, and the audience at WCCFL 33 and AFLA 22 for comments and suggestions on this study. Our sincere thanks to our consultants: Min-ying Sun, Chung-yi Tsai [*Puyuma*], Chin-hua Wu-Tseng, Wan-sung Lin [*Amis*], Yu-ju Chu, Chiu-kui Tai, and Yu-hsia Lin [*Seediq*].

- (1) a. ma-ladram=ku *kan* *Atrung_i* [dra m-uka i Arasip *eci*].¹
 AV-know=1SG.ABS SG.OBL *A_i* [C AV-go LOC A *eci*]
 ‘I know (about Atrung) that she went to Arasip.’ PUYUMA
- b. ma-fana’ kaku *ci* *Lisin_i-an* [Ø tayra i Kalingku *eci*].
 AV-know 1SG.ABS PN *L_i-OBL* [C AV.go LOC K *eci*]
 ‘I know (about Lisin) that she went to Kalingku.’ AMIS
- c. kela-un=mu *ka* *Ikung_i* [Ø m-usa Ø Skangki *eci*].
 know-PV=1SG.ERG ABS *I_i* [C AV-go LOC S *eci*]
 ‘I know (about Ikung) that she went to Skangki.’ SEEDIQ

As shown in the data above, all three languages allow an element that is thematically linked to the Absolutive (ABS) (i.e., pivot, the only argument that can be A'-extracted in a clause) argument of the embedded clause to occur to the left of the embedded predicate, and behave like a matrix object. In (1a)-(b), the “raised phrase” (henceforth XP) appears to receive Oblique (OBL) case from the matrix Actor-voice (AV) verb; in (1c), the XP appears to receive ABS case from the matrix Patient-voice (PV) verb.

Past studies have shown that putative RTO constructions in different languages exhibit considerable cross-linguistic variation with respect to (i) the structure of the embedded clause, (ii) the actual position of the XP, (iii) restrictions on the phrase inside the embedded clause that is co-referential with XP (henceforth YP), and (iv) how the relationship between XP and YP is established (e.g., Massam 1985; Bruening 2001; Polinsky and Potsdam 2001; Alboiu and Hill 2013). As will be shown in this paper, the RTO constructions in these three Formosan languages exhibit similar behaviors in terms of (i) and (ii), yet diverge from one another in terms of (iii) and (iv). Despite the superficial similarities evident in (1), RTO in Amis and Seediq requires the YP inside the embedded clause to be an ABS-marked phrase, while such a restriction is absent in Puyuma RTO. Further, the XP-YP relation is sensitive to islands in Seediq RTO, while the same relation is island-insensitive in Amis and Puyuma RTO. An obvious question is where these variations come from.

In this paper, we argue that the similarities and differences observed among the three languages can be accounted for if they are analyzed as instances of embedded left-dislocation that utilize three different independently motivated strategies to establish the relation between the left-dislocated phrase (the XP) and the CP. Following Landau (2011), we argue that in RTO, a CP and an XP may establish their relation in three different ways. First, the relation between a saturated (i.e., propositional) CP and an XP may be established pragmatically through the *aboutness condition* (2a), which is what we propose to be the case in Puyuma RTO. Alternatively, a CP-XP relation may be established via co-

¹ Abbreviations: ABS=absolutive, AGR=agreement, AGT=agent, AV=Actor voice, C=complementizer, DF=definite, ERG=ergative, ID=indefinite, IRR=irrealis, LK=linker, LOC=locative, LV=Locative voice, NEG=negation, NMZ=nominalizer, OBL=oblique, PRF=perfective, PL=plural, POSS=possessive, PV = Patient voice, SG=singular.

2.1. Similarities

2.1.1. Fully Productive RTO with Finite CP Complements

RTO across the three languages is associated with knowledge/perception verbs (e.g., *know, see, dream, hear, forget, fear*) and finite CP complements. The finite CP analysis for RTO complements is motivated by morphosyntactic differences between RTO and infinitival complements in these languages. Unlike infinitives (e.g., control and restructuring complements) that disallow aspect markers and voice markers other than AV, the embedded complements of RTO in all three languages are fully compatible with different aspect markers and voice markers. Additionally, finite CPs and RTO complements in Puyuma obligatorily carry an overt complementizer, which never appears inside infinitives (Table 2). According to our fieldwork, RTO across the three languages is fully productive and compatible with all CP-taking verbs.

	finite CPs/RTO complements	control/restructuring infinitives
<i>Aspect markers</i>	✓	×
<i>Unrestricted voice type</i>	✓	× (AV-only)
<i>*Overt complementizer</i>	✓ *Puyuma	×

Table 2: Morphosyntactic constraints in types of complementation across the three languages

2.1.2. The Matrix Behavior of the XPs

Across the three languages, XP behaves like a matrix object in terms of case-licensing and reflexive binding. The case status of an XP is always determined by the appropriate matrix case licensor and presents no case-connectedness effect, as can be seen in (3a)-(3c). The XP is thematically identified with the ABS argument of the embedded clause, yet bears OBL case under a matrix AV verb.

(3) The absence of case-connectedness effect in RTO

- a. ma-ladram=ku *kana ngiawi* [dra karatr *ec_i(ABS)* *kana suan*].
 AV-know=1SG.ABS DF.OBL *cat* [C <AV>bite *ec_i(ABS)* DF.OBL *dog*]
 ‘I know (about the cat) that it bit the dog.’ PUYUMA
- b. ma-fana’ kaku *ci Sawmah-ani* [Ø mi-sakilif *ec_i(ABS)* *ci Kulas-an*].
 AV-know 1SG.ABS PN S-OBL [C AV-lie *ec_i(ABS)* PN K-OBL]
 ‘I know (about Sawmah) that she lied to Kulas.’ AMIS
- c. me-’isug=ku *Imin-Ø* [Ø s<m>ipaq *huling=mu ec_i(ABS)*].
 AV-fear=1SG.ABS I-OBL [C <AV>beat *dog=1SG.POSS.(OBL) ec_i(ABS)*]
 ‘I fear (about Imin) that she will beat my dog.’ SEEDIQ

In addition, when XPs are anaphors, they must be bound by the matrix external argument. In RTO sentences with a reflexive XP (4)-(6), the reflexive XP can only be co-indexed with the matrix subject ((4b), (5b), and (6b)) and never

with the embedded subject. As a result, RTO with a reflexive XP always receives a different interpretation from its non-raising counterparts ((4a), (5a), and (6a)).

- (4) A reflexive XP is obligatorily bound by the matrix subject
- a. ma-tiya i Kisaw [dra tu=satra'-ay kan Siber kantaaw].
 AV-dream SG.ABS K [C 3.ERG=slap-LV SG.ERG S himself.ABS]
 'Kisaw_i dreamt that [Siber_k slapped himself_{*i/k}].'
- b. ma-tiya i Kisaw *taytaaw* [dra tu=satra'-ay kan Siber *eci*].
 AV-dream SG.ABS K *himself.OBL* [C 3.ERG=slap-LV SG.ERG S *eci*]
 'Kisaw_i dreamt that [Siber_k slapped him_{i/*k}].' PUYUMA
- (5) A reflexive XP is obligatorily bound by the matrix subject
- a. ma-lemed ni Kulas [Ø ma-palu ni Mayaw cingra].
 PV-dream ERG K [C PV-beat ERG M himself.ABS]
 'Kulas_i dreamt that [Mayaw_k beat himself_{*i/k}].'
- b. ma-lemed ni Kulas *cingra* [Ø ma-palu ni Mayaw *eci*].
 PV-dream ERG K *himself.ABS* [C PV-beat ERG M *eci*]
 'Kulas_i dreamt that [Mayaw_k beat him_{i/*k}].' AMIS
- (6) A reflexive XP is obligatorily bound by the matrix subject
- a. spi-an Ø Walis_i [Ø s<n>ipaq Watan_k ka heya nanak_k].
 dream-LV ERG W [C <PV>beat Watan.ERG ABS himself_k]
 'Walis_i dreamed that Watan_k slapped himself_{*i/k}.'
- b. spi-an Ø Walis_i *ka heya nanak_i* [Ø s<n>ipaq Ø Watan_k *eci*].
 dream-LV ERG W *ABS himself_i* [C <PV>beat ERG W *eci*]
 'Walis_i dreamed that [Watan_k slapped him_{i/*k}].' SEEDIQ

2.1.3. The Absence of Reconstruction Effects

In all three languages, the XP in RTO shows no reconstruction effect. As shown below (7)-(9), when a simple CP complement involves a quantifier subject and a possessive object, a bound variable reading of the possessive object obtains. Therefore, in (7a), the referent of *kantu walak* '3.POSS.OBL child' depends on the referent of its binder *taynaynayan* 'all mothers'. In contrast, if the possessive direct object is "raised" and becomes an XP in RTO (7b), the bound variable reading is no longer available. Hence, the possessive XP can only receive a specific interpretation. Under (7), for instance, the XP can only refer to a specific set of children, i.e., *I know that all mothers_{i/*k} love their_{*i/k} children* (i.e., the children of a certain couple rather than of the mothers).

- (7) The absence of reconstruction effect
- a. ma-ladram=ku [dra sagar kantu walak na taynaynayan].
 AV-know=1SG.ABS [C like.AV 3.POSS.OBL child DF.ABS all.mothers]
 'I know that all mothers_i love their_{i/*k} children.'

- b. ma-ladram=ku *kantu walak* [dra sagar *eci* na taynaynayan].
 AV-know=1SG.ABS 3.POSS.OBL *child* [C like.AV *eci* DF.ABS all.mothers]
 ‘I know that all mothers_i love their_{i/k} children.’

(8) The absence of reconstruction effect

- a. ma-lemed aku [Ø maemin ma-palu nuna wawa ku wacu nira].
 PV-dream 1SG.ERG [C all PV-beat that.ERG child ABS dog 3SG.POSS]
 ‘I dream that all children_i beat their_{i/*k} dogs.’
- b. ma-lemed aku *ku wacu nira* [Ø maemin ma-palu nuna wawa *eci*].
 PV-dream 1SG.ERG ABS *dog* 3SG.POSS [C all PV-beat that.ERG child *eci*]
 ‘I dream that all children_i beat their_{i/k} dogs.’

(9) The absence of reconstruction effect

- a. kela-un=mu [Ø qelu-un bi de-bubu ka laqi=deha].
 know-PV=1SG.ERG [C love-PV very all-mothers ABS child=3PL.POSS]
 ‘I know that all mothers_i love their_{i/*k} children.’
- b. kela-un=mu *ka laqi=deha* [Ø qelu-un bi de-bubu *eci*].
 know-PV=1SG.ERG ABS *child=3PL.POSS* [C love-PV very all-mothers *eci*]
 ‘I know that all mothers_i love their children_{i/k}.’

2.2. Differences

Despite the similarities described above, RTO contractions in the three languages diverge in two important regards: (i) restriction on the YP and (ii) sensitivity to islands.

2.2.1. Amis/Seediq vs. Puyuma: The ABS-only Constraint on the YP

First, Puyuma RTO differs from that in Amis and Seediq in that RTO in the latter two languages requires the XP to be identified with the embedded absolutive argument, while such a restriction is absent in Puyuma RTO. As shown in (10a) and (11a), Puyuma RTO allows its XP to be identified with an embedded ergative or oblique argument as long as it is definite, while in Amis and Seediq RTO ((10b-c) and (11b-c)), an XP identified with any non-ABS element results in ungrammaticality.

(10) YP as an embedded oblique

- a. ma-ladram=ku *i Arasipi* [dra m-uka *eci* i Atrung].
 AV-know=1SG.ABS LOC *A_i* [C AV-go *eci* SG.ABS Atrung]
 ‘I know (about Arasip) that Atrung went there.’
- b. *ma-fana’ kaku *i Kalingku_i* [Ø tayra *eci* Ø-ci Lisin].
 AV-know 1SG.ABS LOC *K_i* [C AV.go *eci* ABS-PN Lisin]
 (‘I know (about Kalingku) that Lisin went there.’)
- c. *kela-un=mu Ø *Skangki_i* [Ø m-usa *eci* ka Ikung].
 know-LV=1SG.ERG LOC *S_i* [C AV-go *eci* ABS Ikung]
 (‘I know (about Skangki) that Ikung went there.’)

PUYUMA

AMIS

SEEDIQ

(11) YP as an embedded ergative

- a. ma-ladram=ku *kan Isawi* [dra tu=trakaw-aw na ngiaw *eci*].
 AV-know=1SG.ABS SG.OBL *I_i* [C 3.ERG=steal-PV DF.ABS cat *eci*]
 ‘I know (about Isaw) that he stole your cat.’ (PUYUMA)
- b. *ma-fana’ kaku *ci Sawmah-ani* [Ø ma-keter *eci* kuna wacu].
 AV-know 1SG.ABS PN *S-OBL_i* [C PV-scold *eci* that.ABS dog]
 (‘I know (about Sawmah) that she scolded that dog.’) (AMIS)
- c. *kela-un=mu *ka Hubi* [Ø ’uq-un *eci* ka sari=su].
 know-PV=1SG.ERG ABS *H_i* [C eat-PV *eci* ABS taro=2SG.POSS]
 (‘I know (about Hubi) that she ate your taro.’) (SEEDIQ)

2.2.2. Seediq vs. Amis/Puyuma: The Island-sensitivity of the XP-YP Relation

Seediq RTO is distinct from RTO in the other two languages regarding its XP-YP relation’s sensitivity to islands. In Seediq, the XP-YP relation respects both complex NP and adjunct islands ((12a) and (13a)), while the same relation in Puyuma and Amis RTO is immune to these islands ((12b-c) and (13b-c)).

(12) Complex NP islands

- a. kilengaw=ku *kan Isawi* [dra ma-ladram=yu [kana kasaerueru
 AV.hear=1SG.ABS SG.OBL *I_i* [C AV-know=2SG.ABS [DF.OBL anecdote
 [dra sagar *eci* dra le’u]]].
 [C like.AV *eci* OBL owl]]] (PUYUMA)
 ‘I heard (about Isaw) the anecdote that he likes owls.’
- b. ma-tengil aku Ø-*ci Kulas* [Ø ma-fana’ kisu [a califacif
 PV-hear 1SG.ERG ABS-PN *K_i* [C AV-know 2SG.ABS [LK rumor
 [Ø ma-ulah *eci* takuwanan]]].
 [C AV-like *eci* 1SG.OBL]]] (AMIS)
 ‘I heard (about Kulas) that you know the rumor that he likes me.’
- c. *q<um>bahang=ku *Hubi*.Ø [Ø kela-un=su [ka kari shelisum
 <AV>hear=1SG.ABS *H_i*-OBL [C know-PV=2SG.ERG [ABS anecdote
 [Ø qiyut *eci* babuy-Ø]]].
 [C <AV>bite *eci* pig-OBL]]] (SEEDIQ)
 (‘I heard (about Hubi) that you know the anecdote that she bit pigs.’)

(13) Adjunct islands

- a. ma-ladram=ku *kan Siber* [dra ka-ilemus=ku
 AV-know=1SG.ABS SG.OBL *S_i* [C IRR.AV-be.angry=1SG.ABS
 [an tu=pukpuk-aw=yu *eci*]].
 [if 3.ERG=beat-PV=2SG.ABS *eci*]]] (PUYUMA)
 ‘I know (about Siber) that I will be angry if he beats you.’

- b. ma-fana' Ø-ci Kulas tuna wacu_i [Ø t<um>angic kaku
 AV-know ABS-PN K OBL.that dog [C <AV>cry 1SG.ABS
 [anu ma-patay ec_i]].
 [if AV-die ec_i]]

AMIS

‘Kulas knows (about the dog) that I will cry if it dies.’

- c. *kela-un=mu ka Imin_i [ya'asa m-huqil ka huling
 know-PV=1SG.ERG ABS I_i [because AV-die ABS dog
 [kika lingis ec_i]].
 [so <AV>cry ec_i]]

SEEDIQ

(‘I heard (about Imin) that because the dog died, she is crying.)

2.3. Section Summary

In sum, RTO constructions in Puyuma, Amis, and Seediq are similar in terms of (i) the finite CP status of their complements, and (ii) the lack of reconstruction effects with their XPs, while they diverge in terms of the case restriction on the YP and the XP-YP relation’s sensitivity to islands. Puyuma RTO is the most unrestricted, imposing no restriction on YP or the XP-YP relation. Amis RTO obeys the absolutive-only constraint on the YP, but shows no island-sensitivity of the XP-YP relation. Seediq RTO is the most restricted, as it is subject to both the absolutive-only constraint on the YP and island conditions. Table 3 summarizes the empirical observations presented so far.

	PUYUMA	AMIS	SEEDIQ
<i>Reconstruction effects</i>	×	×	×
<i>ABS-restriction on the YP</i>	×	✓	✓
<i>Island-sensitivity</i>	×	×	✓

Table 3: Similarities and differences in RTO across the three languages

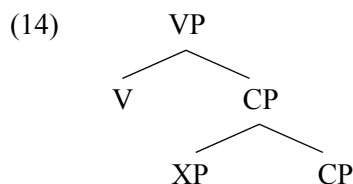
3. Formosan RTO as a Case of Embedded Left-Dislocation

Based on the empirical observations discussed above, we present our analysis of the basic structure of RTO, according to which RTO constructions are instances of embedded left-dislocation. We first motivate the left-dislocated adjunct analysis for the XPs based on the similarities between XPs and base-generated left-dislocated phrases. We then present our account of the matrix object-like behavior of XPs.

3.1. Formosan RTO Involves a Finite CP with a Base-Generated Left-Dislocated XP

In the previous discussion (section 2.1.1), we showed that RTO in the three languages involves a finite CP complement, based on the unrestricted aspect and voice marking and overt complementizer observed in the RTO complements. In terms of the

structural relationship between the XP and the CP, we propose that the XPs are base-generated left-dislocated phrases that adjoin to the embedded CPs, as in (14).



This analysis is motivated by the observation that XPs in Formosan RTO share a number of similarities with base-generated left-dislocated phrases cross-linguistically, such as *left-dislocated phrases* in Niuean (Massam 1985) and Italian (Cinque 1990), *external topics* (as opposed to internal topics) in Mayan languages (Aissen 1992), and *hanging topics* in Greek (Anagnostopoulou 1997). As discussed earlier, XPs in Formosan RTO (i) lack reconstruction effects (2.1.3), and (ii) present no case connectedness effects (2.1.2). Both characteristics are typical of hanging/external topics cross-linguistically. RTO in Puyuma further provides an additional motivation for the base-generated left-dislocated adjunct analysis of XPs. As shown in section 2.2, Puyuma RTO imposes no case restriction on the YP, nor does it exhibit sensitivity to islands. At the same time, however, it must satisfy the *aboutness condition*, i.e., the content of the CP must be “about” the referent of the XP, a constraint that is commonly observed between base-generated left-dislocated phrases and the clauses that follow them. As can be seen in the following example, the failure to satisfy the aboutness condition results in unacceptability in Puyuma RTO (15).

- (15) *ma-tiya=ku kan Siber [dra trakaw dra paysu i Isaw].
 AV-dream=1SG.ABS SG.OBL S [C <AV>steal ID.OBL money SG.ABS I]
 (‘I dreamt (about Siber) that Isaw stole money.’)

PUYUMA

Based on these observations, we argue that XPs are best analyzed as base-generated left-dislocated adjuncts to the CP.

3.2. The Matrix Behavior of XPs

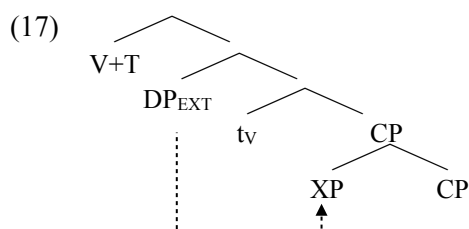
As discussed in section 2.1, XPs in Formosan RTO constructions exhibit typical characteristics of matrix objects in terms of case marking and anaphora binding. In this subsection, we provide our accounts for these behaviors based on the proposed analysis of XPs as embedded-dislocated phrases.

First, as already shown in (7)-(9), when XPs are reflexive pronouns, they are obligatorily bound by the matrix subjects. We argue that the reflexivization facts follow straightforwardly from the proposed structure of RTO in (14), according to which the XP is the closest DP c-commanded by the matrix subject. For the sake of

concreteness, we adopt the A-chain analysis for reflexives proposed by Reinhart and Reuland (1993), which is based on the following definition of an A-chain.

- (16) *An A-chain, under its broadest definition, is any sequence of coindexation that is headed by an A-position and satisfies antecedent government; that is, each coindexed link, except for the head, is c-commanded (i.e., m-commanded) by another link, and there is no barrier between any two of the links. (693).*

According to (16), if an XP is pronominal, it must be reflexivized, as it is the closest DP c-commanded by the matrix subject, i.e., the head of the A-chain.



Second, as can be seen in all RTO data in this paper, an XP must bear the morphological case that would be borne by a direct object. We argue that such case marking of XPs follows from two independently motivated generalizations of case in Formosan languages. First, CPs need to be “case-licensed” in Formosan languages, *i.e.*, *CPs enter into an Agree relation with a case-licensing head in these languages.* Second, as mentioned earlier, *there must be one and only one ABS-marked phrase in a given clause.* As will be discussed below, these two generalizations lead us to propose that the morphological case borne by XPs is a reflex of the case assigned to the co-occurring CP. Namely, the XP inherits the case that is assigned to the XP-CP constituent.

Empirical motivations for the first generalization come from the A'-extraction asymmetry that has been reported in several Philippine-type Austronesian languages including Chamorro (Chung 1991, 1994, 1998), Tagalog (Rackowski and Richards 2005), and Malagasy (Pearson 2009). According to Chung, in Chamorro, the extraction of a *wh*-phrase is accompanied by voice agreement (“*wh*-agreement”) on the verb, which indicates the case status of the *wh*-phrase. In (18), the embedded verb has an infix <*in*>, which indicates that the *wh*-phrase extracted bears OBL case. In the same example, the matrix verb has a prefix *um-*, which corresponds to the nominative case. As the only element that could bear nominative case in (18) is the CP, the example motivates the argument that Chamorro verbs agree with CPs in terms of case.

- (18) *Hayi*_i um-istotba si-Juan [*ec*_i ni m<*in*>ahalang i asagua-ña *ec*_i]?
*who*_i NOM-disturb PN-J [*ec*_i C OBL.lonely the spouse-AGR *ec*_i]
 ‘Who does it disturb Juan that his wife is lonely for?’ (Chung 1994:14)

CHAMORRO

A similar proposal has been put forward for Tagalog (e.g., Rackowski and Richards 2005; Law 2014), as exemplified in the following Tagalog relative clause data:

- (19) Long-distance A'-extraction in Tagalog
- a. Gusto ko ang *libro_i*-ng [s<in>abi ni Fred [na b<in>ili ni Maria *ec_i*]].
 like 1SG ABS *book_i*-LK [<PV.PRF>say ERG F [LK <PV.PRF>buy ERGM *ec_i*]]
 'I like the book that Fred said that Maria bought.'
- b. *Gusto ko ang *libro_i*-ng [nag-sabi si Fred [na b<in>ili ni Maria *ec_i*]].
 like 1SG ABS *book_i*-LK [AV.PRF-say ABS F [LK <PV.PRF>buy ERG M *ec_i*]]
 ('I like the book that Fred said that Maria bought.')(Law 2014:4; glosses ours)

In (19a-b), the head of the relative clause *libro* 'book' is A'-extracted from the most deeply embedded clause. As is well known, A'-extractions in Tagalog impose the absolutive-only condition, i.e., only ABS-marked phrases can be A'-extracted. However, while *libro* 'book' in both (19a) and (19b) is ABS-marked inside the most deeply embedded clause, only (19a) is grammatical. According to Rackowski and Richards (2005) and Law (2014), the crucial difference between the two lies in the fact that the voice of the intermediate verb *sabi* 'say' is PV in (19a), but AV in (19b). Hence, the most deeply embedded CP in the sentences in (19) can be analyzed as ABS-licensed by the PV-form of the verb *sabi* in (19a) but OBL-licensed by the AV-form of the same verb in (19b). This extraction asymmetry led to the following generalization:

- (20) *Only those CPs and DPs that Agree with a phase head on independent grounds (e.g., direct objects and complement clauses) are transparent for wh-extraction.* (Rackowski and Richards 2005:582)

We argue that the same generalization applies to the three Formosan languages, although its effects are manifested slightly differently. As in Tagalog, A'-extractions in Puyuma, Amis, and Seediq obey the ABS-only constraint. Consider the following *wh*-questions from Amis.³

- (21) Long-distance A'-extraction in Amis
- a. *cima_i* ku [ka-tengil-an isu [ku fangcal-ay *ec_i*]]?
who_i ABS [PV.IRR-hear-LV.NMZ 2SG.ERG [ABS pretty-AGT.NMZ *ec_i*]]
 'Who do you hear is pretty?'
- b. *cima_i* ku [ka-fana'*(-an) ✓isu/*kisu [ku fangcal-ay *ec_i*]]?
who_i ABS [AV.IRR-know-*(LV.NMZ) ✓2SG.ERG/*ABS [ABS pretty-AGT.NMZ *ec_i*]]
 'Who do you know is pretty?'

In (21a), the intermediate embedded clause has the verb *tangil* 'hear' in non-AV form, which presumably makes the most deeply embedded CP ABS-licensed,

³ The same pattern is observed in Puyuma and Seediq. Due to space limitations, only the Amis relevant data are not presented here.

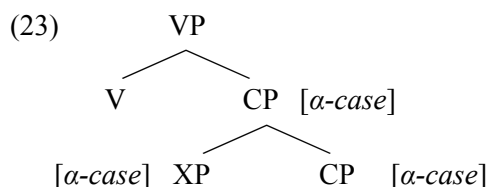
similar to the Tagalog example (19a). As expected, the sentence is grammatical. (21b), on the other hand, involves the verb *fana* ‘know’ as the verb of the intermediate embedded clause. Under normal circumstances, the verb *fana* is only compatible with AV structure, as in (22).

- (22) ✓ ma-fana’/*fana’-en/*-an kisu/*isu [Ø ma-ulah kaku ci Lisin-an].
 ✓ AV-know/*know-PV/*-LV 2SG.ABS/*2SG.ERG [C AV-like 1SG.ABS PN L-OBL]
 ‘You know that I like Lisin.’

AMIS

However, despite the fact that the verb *fana* ‘know’ normally takes an AV marker (22), in A’-extraction sentences like (21b), the same verb is observed to obligatorily co-occur with a non-AV nominalizer -an, indicating that the verb is functioning as a non-AV verb under extraction circumstances. We believe that this exceptional behavior of the verb under A’-extractions is due to the generalization in (20). Namely, if the verb behaved as it normally does, the most deeply embedded CP would be OBL-marked, hence an A’-extraction out of this CP would be impossible. Only if the verb in (21b) is non-AV-marked would the most deeply embedded CP receive ABS case and allow A’-extractions.⁴ This, in turn, supports the claim that CPs are “case-licensed” in Formosan languages.

Going on from this generalization, we propose that (i) in Formosan RTO, XPs always bear the case assigned to the co-occurring CP complement, and (ii) the XP and CP must form a constituent. Following these two proposals, we argue that the case-marking on the XP is a realization of the case assigned to the XP-CP constituent, as schematized in (23).



Supporting evidence for proposal (ii) come from the double-Absolutive-marking observed in Puyuma RTO with complex DP islands, along with the *Absolutive-as-unique* condition shared by the three languages. As shown in the following Puyuma data, when an XP is adjoined to a complex DP, the XP and the complex DP bear the same morphological case, OBL and ABS, respectively, as conditioned by the matrix voice type.

⁴ Importantly, the *ka*-prefix in (21a) serves as a PV marker, while that in (21b) serves as an AV marker, as evidenced by the different case statuses (ERG or ABS) of the external arguments introduced by the verbs in simple clauses.

- (24) RTO with an OBL-marked complex DP island in Puyuma
 aparu=ku $\left(\begin{array}{c} \textit{kan} \textit{ Isaw} \left[\textit{DP kana ngay} \right] \left[\textit{dra salem ec}_i \textit{ dra bujir} \right] \right] \\ \textit{forget.AV=1SG.ABS} \left[\textit{SG.OBL I}_i \left[\textit{DP DF.OBL rumor} \left[\textit{C <AV>grow ec}_i \textit{ ID.OBL taro} \right] \right] \right] \end{array} \right)$
 ‘I forgot (about Isaw) the rumor that he grows taros.’
- (25) RTO with an ABS-marked complex DP island in Puyuma
 ku=aparu-ay $\left(\begin{array}{c} \textit{i} \textit{ Isaw} \left[\textit{DP na ngay} \right] \left[\textit{dra salem ec}_i \textit{ dra bujir} \right] \right] \\ \textit{1SG.ERG=forget-LV} \left[\textit{SG.ABS I}_i \left[\textit{DP DF.ABS rumor} \left[\textit{C <AV>grow ec}_i \textit{ ID.OBL taro} \right] \right] \right] \end{array} \right)$
 ‘I forgot (about Isaw) the rumor that he grows taros.’
*The NP ‘rumor’ can either precede or follow the CP.

In (24)-(25), the XP ‘*Isaw*’ is adjoined to the complex DP *the rumor [that he grows mangos]*. Importantly, in (25), both the XP *Isaw* and the head of the DP *ngay* ‘rumor’ bear ABS case under the matrix verb *aparu* ‘forget’ that is in non-AV form. According to the *Absolutive-as-unique* generalization which states that *there can only be one ABS-marked phrase in a given clause*, the XP *Isaw* and the DP *the rumor [that he grows mangos]* can only be analyzed as sharing the same Absolutive case assigned by the matrix clause. We argue that essentially the same process takes place in RTO, with an XP adjoined to a simple CP complement. The only difference between the situation of the complex DP and that of the CP complement is that the (abstract) case-status does not have morphological realization on CP complements.

This motivates the analysis that an XP and a CP in RTO must form a single constituent and share the same Case. That is, the case borne by the XP could *not* have been independently assigned to it, as such an analysis would violate the *Absolutive-as-unique* generalization. Thus, we propose that the case realized on the XPs in RTO is a reflex of the case assigned to the XP-CP constituent, as schematized in (23) above.

4. Three Strategies to Establish the XP-YP Relation

Building on the analysis of the structure of RTO proposed above, in this section, we present our proposal for the micro-variation in RTO constructions in the three languages. We argue that the micro-variation derives from three independently motivated strategies that the three languages utilize to establish the relation between the left-dislocated phrases (XPs) and CPs, as in (26).

- (26) a. $V_{\text{MATRIX}} \text{XP} + \left[\text{CP} \right]$ via aboutness condition [Puyuma RTO]
 b. $V_{\text{MATRIX}} \text{XP}_i \left[\text{CP Op}_i \textit{ pronoun}_i \right]$ [Amis RTO]
 c. $V_{\text{MATRIX}} \text{XP}_i \left[\text{CP Op}_i \textit{ <t}_i \textit{>} \right]$ [Seediq RTO]

4.1. Puyuma RTO: Propositional CP and the Aboutness Condition

As shown in section 2, Puyuma RTO is distinct from the RTO in Amis and Seediq in that the former imposes no syntactic restriction on the relation between XPs and

CPs. Further, the XP-YP relation in Puyuma RTO is immune to islands. We propose that this is because the Puyuma RTO involves XPs with propositional CP complements, whose relationship is established pragmatically through the aboutness condition. In other words, the only requirement imposed on the XP-CP relation is that the content of CPs be “about” the referents of XPs.

(27) $V_{\text{MATRIX}} \quad XP_{(i)} \quad [CP \quad (YP_i) \quad]$

Following previous studies, we assume the aboutness condition is what licenses the instances of left-dislocation-like phenomena which involve no gap/pronominal copy, such as *external topics* in Mayan (Aissen 1991), *hanging topics* in Greek (Anagnostopoulou 1997), “*major subjects*” in Japanese (Saito 1982; Kuroda 1986; Heycock 1994, among others), and *copy-raising without a pronoun* in English and Hebrew (Landau 2011). The aboutness condition is often satisfied via co-reference between an XP and a phrase inside the CP (YP), yet such co-indexization is not necessary. Therefore, Puyuma RTO is felicitous even when the XP does not refer to an argument of the embedded clause (28)-(29).

(28) kilegnaw=ku kana sasaya na bulran [dra pi<a>walak i Pilay].
 AV.hear=1SG.ABS DF.OBL one LK month [C <IRR>give.birth SG.ABS P]
 ‘I heard that Pilay is going to give birth to her child next month.’ (PUYUMA)

(29) ma-ladram=ku an milanang dra bira’ [dra u<a>ruma=yu].
 AV-know=1SG.ABS when AV.be.yellow ID.OBL leave [C <IRR>AV.be.back=2SG.ABS]
 ‘I know that you will be back when the leaves turn yellow.’ (PUYUMA)

4.2. Amis RTO: Predicative CP with a Base-Generated Null Topic Op in [Spec, CP]

Unlike Puyuma RTO, Amis RTO imposes the ABS-only requirement on YPs; namely, XPs must be co-indexed with an ABS-marked argument inside the CP ((10b) and (11b)). On the other hand, unlike Seediq RTO and like Puyuma RTO, the XP-YP relation in Amis RTO can cross islands ((12b) and (13b)).

We argue that the two characteristics observed in Amis can be captured under the analysis that Amis RTO involves a predicative CP that involves a base-generated null operator in [Spec, CP]. Importantly, this base-generated null operator is a *topic* that must be co-referential with an ABS-marked pronoun in the embedded clause, which can be either covert or overt. This null topic operator, in turn, is co-indexed with the XP (30).

(30) $V_{\text{MATRIX}} \quad XP_i \quad [CP \quad Op_i \quad (ABS\text{-pronoun}_i) \quad]$

According to Landau (2011), the configuration in (30) accounts for the cases of *left-dislocation* and *copy-raising* that require the presence of a pronoun, as well as *prolepsis* constructions in languages like Madurese (Davies 2005). Under this

analysis, the properties of Amis RTO follow straightforwardly from the properties of topic constructions in Amis. According to our data, while topics in Amis must be co-indexed with an ABS-marked phrase, this co-reference relation does not respect syntactic locality or islands. In (31a) below, the topic *Ofad* is co-referential with the ABS-marked pronoun *cingra* ‘he’ embedded inside the adverbial clause (which can be either null or overt). As can be seen, this co-reference “skips” the closer ABS-marked phrase, *kaku* ‘I’, and also crosses an adjunct island. In (31b), the topic *Mayaw* is co-referential with the ABS-marked pronoun *cingra* ‘he’ inside the complex DP *sinpung* ‘news’.

(31) The absence of locality condition in Amis topic constructions

- a. \emptyset -ci *Ofad*_i, tayra kaku i Busung,
 ABS-PN *O*_i AV.go 1SG.ABS LOC B
 [anu pafli (*cingra*_i) takuwanan tu paysu].
 [if AV.give (3SG.ABS) 1SG.ABS OBL money]
 ‘Ofad_i, I will go to Busung if he_i gives me money.’

- b. \emptyset -ci *Mayaw*_i, ma-tengil aku
 ABS-PN *M*_i PV-hear 1SG.ERG
 [kuna sinpung [adada (*cingra*_i)]].
 [ABS.that news [AV.be.sick (3sg.abs)]]
 ‘Mayaw_i, I heard the news that he_i is sick.’

Based on the parallel behavior between XPs and topics in Amis, we propose that there is a null topic operator inside the CP in Amis RTO (32). This accounts for the ABS-only requirement imposed on YPs, as well as the lack of island-sensitivity observed in the same construction.

- (32) ma-fana’ kaku *ci* *Mayaw-an*_i [kuna sinpung [*Op*_i adada (*cingra*_i)]].
 AV-know 1SG.ABS PN M-OBL_i [ABS.that news [*Op*_i AV.be.sick (3sg.abs)]]
 ‘I know (about Mayaw_i) the news that he_i is sick.’

AMIS

4.3. Seediq RTO: Predicative CP with an A’-moved Null Operator in [Spec, CP]

Among the three languages’ RTO constructions, Seediq RTO is unique in that its XP-YP relation is island-sensitive ((11c) and (12c)). Further, it imposes the ABS-only requirement on YPs ((10b) and (11b)), as also observed in Amis. We argue that Seediq RTO exhibits these characteristics because it involves a predicative CP and a null operator that A’-moves from the theta-position to [Spec, CP], as in (33).

- (33) V_{MATRIX} XP_i [CP *Op*_i *t*_i]

The proposed structure in (33) offers a unified account for the ABS-only requirement and the island-sensitivity in Seediq RTO. As the construction involves

an A'-movement, (i) the YP must be an ABS-marked phrase, as only an ABS-marked phrase can be A'-extracted, and (ii) the XP-YP relation cannot cross islands, as an A'-trace cannot be licensed with islands.

5. Conclusions and Implications

We have analyzed RTO in the three Formosan languages as instances of embedded left-dislocation that involve a finite CP and a base-generated XP as a left-dislocated adjunct. We have also argued that the micro-variation observed in the three languages can be captured under the hypothesis that a CP can establish a relation with a left-dislocated phrase in three different ways: (i) via the *aboutness condition*, (ii) via a *base-generated null operator*, and (iii) via an *A'-moved null operator*. If the present analysis is on the right track, the RTO constructions in the three Formosan languages provide novel support for Landau's (2011) claim that *a CP and a CP-external phrase may establish a connection via three different strategies*.

The present proposal adds RTOs in Puyuma, Amis, and Seediq to the growing list of languages with RTO-like phenomena that do not involve movement out of finite CPs, including Niuean and Fijian RTO (Massam 1985), Malagasy RTO (Paul and Rabaovololona 1998; Pearson 2005), Tsez long-distance agreement (Polinsky and Potsdam 2001), Madurese prolepsis (Davies 2005), and Sundanese prolepsis (Kurniawan 2011). What is conspicuously missing in this picture of Formosan RTO is a strategy that involves actual syntactic movement of XPs out of CPs. In fact, only a handful of languages with RTO-like phenomena have been argued to involve syntactic movement of XPs out of finite CPs (e.g., Japanese: Kuno 1976 and Tanaka 2002; Korean: Yoon 2007; Romanian: Alboiu and Hill 2013; and Zulu: Halpert and Zeller to appear). The rareness of instances of actual raising out of finite CPs in typologically and geographically diverse languages suggests the pervasiveness of the generalization that *A-movement out of finite CPs is prohibited* (e.g., Tensed-S Condition; Chomsky 1973), and seems to be consistent with the intuition that *finite CPs form an independent and complete unit with respect to syntactic operations* (Phase Impenetrability Condition; Chomsky 2000, 2001, 2008). The existence of the languages that arguably do allow A-movement out of finite CPs, on the other hand, raises the question of under what circumstances a language may allow such an operation.⁵

⁵ According to Bruening (2001), Passamaquoddy RTO may be added to this list, as it arguably utilizes two different strategies in establishing the XP-YP relation, one of which involves an XP base-generated at the left edge of the CP, and undergoing A-movement to the matrix domain.

References

- Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68:43-80.
- Alboiu, Gabriela & Virginia Hill. 2013. The case of A-bar ECM: Evidence from Romanian. In Stefan Keine & Shayne Sloggett (eds.), *Proceedings of the Forty-Second Annual Meeting of the North East Linguistic Society: 25-39*. Amherst, MA: GSLA.
- Anagnostopoulou, Elena. 1997. Clitic left dislocation and contrastive left dislocation. In Elena Anagnostopoulou, Henk van Riemsdijk, & Frans Zwats (eds.), *Materials on left dislocation*, 151-192. John Benjamins: Amsterdam.
- Blust, Robert. 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun & Paul Li (eds.), *Selected papers from the Eighth International Conference on Austronesian Linguistics*, 31-94. Taipei: Institute of Linguistics, Academia Sinica.
- Bruening, Benjamin. 2001. Raising to object and proper movement. Ms., University of Delaware.
- Chomsky, Noam. 1977. On wh-movement. In Peter W. Culicover, Thomas Wasow, & Adrian Akmajian (eds.), *Formal syntax*, 71-132. New York: Academic Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89-156. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (eds.), Ken Hale: *A life in language*, 1-52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria L. Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133-166. Cambridge, MA: MIT Press.
- Chung, Sandra. 1991. Sentential subjects and proper government in Chamorro. In Carol Geogopoulos & Roberta Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, 77-99. Amsterdam: Kluwer Academic Publishers.
- Chung, Sandra. 1994. Wh-agreement and “referentiality” in Chamorro. *Linguistic Inquiry* 25:1-44.
- Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. Chicago and London: University of Chicago Press.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, MA: MIT Press.
- Davies, William. 2005. Madurese prolepsis and its implications for a typology of raising. *Language* 81, 645-665.
- Halpert, Claire & Jochen Zeller. To appear. Right dislocation and raising-to-object in Zulu. *The Linguistic Review*.

- Heycock, Caroline. 1993. Syntactic predication in Japanese. *Journal of East Asian Linguistics* 2:167-211.
- Kuno, Susumu. 1976. Subject raising. In Masayoshi Shibatani (ed.), *Syntax and semantics 5: Japanese generative grammar*, 17-50. New York: Academic Press.
- Kurniawan, Eri. 2011. Does Sundanese have prolepsis and/or raising-to-object constructions? *Proceedings of the 18th Annual Meeting of the Austronesian Formal Linguistics Society (AFLA 18)*, 16-30. Online publication hosted by the University of Western Ontario.
- Kuroda, Sige-Yuki. 1986. Movement of noun phrases in Japanese. In Takashi Imai & Mamoru Saito (eds.), *Issues in Japanese linguistics*, 229-272. Dordrecht: Foris.
- Landau, Idan. 2011. Predication vs. aboutness in copy raising. *Natural Language and Linguistic Theory* 29:779-813.
- Law, Paul. 2014. The syntax of Tagalog relative clauses. Paper presented at the 21st Annual Meeting of the Austronesian Formal Linguistics Society (AFLA 21). Honolulu, May 23-25, 2014.
- Massam, Diane. 1985. Case theory and the projection principle. Ph.D dissertation, Massachusetts Institute of Technology.
- Paul, Ileana & Lucie Rabaovololona 1998. Raising to object in Malagasy. In *The structure of Malagasy, vol. 2. UCLA Occasional Papers in Linguistics* 20:50-64.
- Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory* 23(2), 381-457.
- Pearson, Matthew. 2009. Another look at no: Pseudo-clefts and temporal clauses in Malagasy. In Sandy Chung, Daniel Finer, Ileana Paul, and Eric Potsdam (eds.), *Proceedings of the Sixteenth Meeting of the Austronesian Formal Linguistics Association*:165-179. University of California, Santa Cruz.
- Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19, 583-646.
- Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565-599.
- Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657-720.
- Saito, Mamoru. 1982. Case marking in Japanese: A preliminary study. Ms., MIT, Cambridge, MA.
- Tanaka, Hidekazu. 2002. Raising to object out of CP. *Linguistic Inquiry* 33:637-652.
- Yoon, James. 2007. Raising of major arguments in Korean (and Japanese). *Natural Language and Linguistic Theory* 25(3):615-653.

VOICE AND REGISTER DIFFERENCES IN MADURESE*

Helen Jeoung
University of Pennsylvania
jeoung@sas.upenn.edu

In this paper I examine the voice system of Madurese, and demonstrate that the speech levels differ with respect to the number of available voices. The polite register of Madurese has an active voice, a canonical passive voice and an object voice. In contrast, the familiar register has only two voices, active and passive. Another grammatical difference between the registers is demonstrated in DP extraction patterns. In polite speech, subject DPs may be extracted, while object DPs may only be extracted when there is null voice morphology on the verb. In familiar speech, subject DPs may be extracted, while object DP extraction is not possible at all. The availability of voices is attributed to a different inventory of Voice⁰ heads in each register. The difference in DP extraction falls out from the lack of object voice and null voice marking in the familiar register.

1. Introduction

In Madurese there are three available voices, which are defined by morphosyntactic criteria. The three-voice distinction is illustrated below in the polite register: active voice (1), passive voice (2) and object voice (3):

- (1) Ramah nembhalih potraepon.¹
father.Pol AV-call.Pol-Appl son.Pol-Def.Pol
'Father called his son.'
- (2) Potraepon etembhalih.
son.Pol-Def.Pol PV-call.Pol-Appl
'The son was called.'
- (3) Potraepon ampon ramah tembhalih.
son.Pol-Def.Pol Perf.Pol father.Pol OV-call.Pol-Appl
'Father called his son.'

* I thank my Madurese consultants, Isya Mahfud and Maimuna, for their help and patience. Thanks also to Julie Legate for many comments and suggestions on this work.

¹ Abbreviations: Appl = applicative; AV = active voice; Def = definite; DP = determiner phrase; Fam = familiar; Neg = negation; OV = object voice; Perf = perfective; Pol = polite register; PP = prepositional phrase; PV = passive voice; Rel = relative marker. Morphemes used in polite register are glossed as polite; familiar register is not specifically indicated.

The goals of this paper are three-fold. First, I refine previous accounts of Madurese morphosyntax (Killiaan 1897; Stevens 1968; Davies 2010) which describe only two voices in the language. I demonstrate that three voices exist in the polite register of Madurese, while in the familiar register only two voices are available. Second, I show that the difference between the registers also extends to DP extraction patterns: polite Madurese allows both subjects and objects to extract, while familiar Madurese allows only subjects to extract. Third, I consider an analysis which can account for the grammatical differences between the registers.

2. Three Voices in Madurese

2.1. Background on Madurese

Madurese (mad) is a Malayo-Polynesian language spoken primarily in western Indonesia, in the province of East Java. Madurese has several dialects; all data provided here is from the western variety of Madurese that is associated with the Bangkalan regency and also spoken in Surabaya and other parts of East Java.

Across all varieties, Madurese is an SVO language, with tense and aspect morphemes occurring between subject and verb. Madurese also has (at least) three registers or speech levels (Davies 1999, 2010; Pawitra 2008; Safioedin 1977; Stevens 1968). The familiar register (*kasar*) is used among close friends and for children, while the polite register (*aloh*) is used in formal settings and to convey respect toward adults. The middle register (*tengnga'an*) is less productive than familiar and polite speech, with fewer dedicated lexical items. Thus middle speech is often mixed with either familiar or polite speech (Davies 2010). For the sake of clarity, this paper examines only the polite and familiar speech levels of Madurese.

2.2. Active Voice (AV)

In the active voice, the Agent occurs in the preverbal position of surface subjects. The verb bears either a nasal prefix *N-* or the prefix *a-*,² which are selected by the verb root:

(4) Polite AV

- a. Ramah nembhalih potraepon.
ramah N-(t)embhal-ih potra-epon
father.Pol AV-call.Pol-Appl son.Pol-Def.Pol
'Father called his son.'
- b. Ramah lo' aparengaghih buku panekah dha' kancaepon.
ramah lo' a-pareng-aghih buku panekah dha' kanca-epon
father.Pol Neg AV-give.Pol-Appl book that.Pol to friend-Def.Pol
'Father did not give the book to his friend.'

² The distribution of *N-* and *a-* is described in Davies (2010). Generally, intransitive verbs take *a-* but there are numerous exceptions, as in (4b).

- (5) Familiar AV
- a. Ali ngatoeh ana'eng.
 Ali N-kato-ih ana'-eng
 Ali AV-call-Appl child-Def
 'Ali called his child.'
- b. Ali lo' aberri' buku jiyah dha' kancanah.
 Ali lo' a-berri' buku jiyah dha' kanca-nah
 Ali Neg AV-give book that to friend-Def
 'Ali did not give the book to his friend.'

The morphology of the voice prefixes does not vary between registers; the AV prefixes *N-* and *a-* are employed in both polite and familiar Madurese.

2.3. Passive Voice (PV)

The passive voice in Madurese is a canonical passive, defined by the following typological characteristics (Haspelmath 2001, Kulikov 2010). An argument other than the Agent occurs as surface subject (typically the Patient/Theme). The Agent may be implicit, and is therefore optional. When the Agent does occur, it is as an oblique argument, embedded in a PP by-phrase.

In related Indonesian languages, the passive voice is marked with a verbal affix (e.g. Indonesian *di-*, Balinese *ka-* or *-a*). Likewise, Madurese passive verbs occur with the prefix *e-*:

- (6) Polite PV
- Potraepon etembhalih (sareng ramah).
 potra-epon e-tembhal-ih sareng ramah
 son.Pol-Def.Pol PV-call.Pol-Appl by.Pol father.Pol
 'The son was called (by father).'
- (7) Familiar PV
- Ana'eng ekatoeh (bi' Ali).
 ana'-eng e-kato-ih bi' Ali
 child-Def PV-call-Appl by Ali
 'The child was called (by Ali).'

The thematic Patient occurs in the preverbal position of subjects in (6) and (7), while the PP expressing the Agent is optional. Just as in AV, the PV prefix does not vary between the registers.

Note that voice prefixes are obligatory for all basic active and passive clauses.³ In contrast to the examples in (4) through (7), the same clauses are ungrammatical with a bare verb:

- (8) Polite
- a. *Ramah tembhalih potraepon. (Active word order)
ramah tembhal-ih potra-epon
father.Polcall.Pol-Appl son.Pol-Def.Pol
- b. *Potraepon tembhalih (sareng ramah). (Passive word order)
potra-epon tembhal-ih sareng ramah
son.Pol-Def.Pol PV-call.Pol-Appl by.Pol father.Pol
- (9) Familiar
- a. *Ali katoeh ana'eng. (Active word order)
Ali kato-ih ana'-eng
Ali call-Appl child-Def
- b. *Ana'eng kato-eh bi' Pak Ali. (Passive word order)
ana'-eng kato-ih bi' Pak Ali
child-Def call-Appl by Mr Ali

Thus “bare” verbs (i.e. verbs lacking overt voice morphology) are not well-formed in basic AV or PV clauses.

2.4. Object Voice (OV)

I now turn to the third voice in Madurese, an object voice⁴ which has not been previously noted in the literature. To identify OV clauses, I follow analyses of object voice in related Indonesian languages, including Arka and Manning 2008, Chung 1976, Cole et al 2006, Cole et al 2008, Guilfoyle et al 1992 for Indonesian; Legate 2012, 2014 for Acehnese; and Arka 2003 for Balinese. Building on this previous body of research, I employ morphosyntactic criteria in differentiating object voice from passive and active voices.

I first examine object voice clauses in the polite register. In object voice, the thematic Patient occurs as grammatical subject, as it does in passive clauses. Unlike PV verbs however, the OV verb is bare, will null voice marking. Other affixes (causative or applicative morphemes) may occur on an OV verb, but the voice affix must be null as in (10):

³ Except for several intransitive verbs like *mole* 'go home,' which never take a voice prefix. Imperatives also employ the bare verb stem.

⁴ The terminology varies in the literature for passive and object voices, respectively: for example, Chung 1976 uses “canonical passive” and “object preposing passive”; Dardjowidjojo 1978 and Sneddon et al 2012 use “Passive type I” and “Passive type II”; Arka and Manning 2008 use “passive voice” and “objective voice”; Aldridge 2008 uses “pronominal passive” for object voice. The Indonesian term *pasif semu* also refers to object voice.

- (10) Polite OV
 Potraepon ampon (*sareng) ramah tembhalih.
 potra-epon ampon sareng ramah Ø-tembhal-ih
 son.Pol-Def.Pol Perf.Pol by.Pol father.Pol OV-call.Pol-Appl
 ‘Father called his son.’

The object voice furthermore differs from the canonical passive voice with regard to the Agent. In the canonical passive, the Agent is not a core argument of the verb, but occurs as an oblique argument, embedded in a PP by-phrase. In object voice however, the Agent is not demoted to an oblique argument; it remains a core argument of the verb. Note that in (10) the Agent *ramah* ‘father’ may not occur with the preposition *sareng* ‘by.’⁵ The OV Agent occurs as a DP rather than a PP.

As already mentioned, in a canonical passive, the presence of the Agent is optional. By contrast, the object voice Agent must always be present, and may not be null. Note that in both AV and PV, Madurese allows subjects and objects to be null (*pro*-dropped) where licensed by the discourse (Davies 2010). Thus the subject in (11) may be null if previously mentioned:

- (11) Polite AV
 ... nembhalih potraepon.
 N-(t)embhal-ih potra-epon
 AV-call.Pol-Appl son.Pol-Def.Pol
 ‘...(he) called his son.’

However, the Agent may not be null in the corresponding OV clause, but must be pronounced:

- (12) Polite OV
 *... potraepon ampon tembhalih.
 potra-epon ampon Ø-tembhal-ih
 son.Pol-Def.Pol Perf.Pol OV-call.Pol-Appl
 ‘...(he) called his son.’

Not only must the Agent be pronounced in an object voice clause, it must also remain in its thematic position, immediately to the left of the verb. Other elements may not intervene between the Agent and verb:

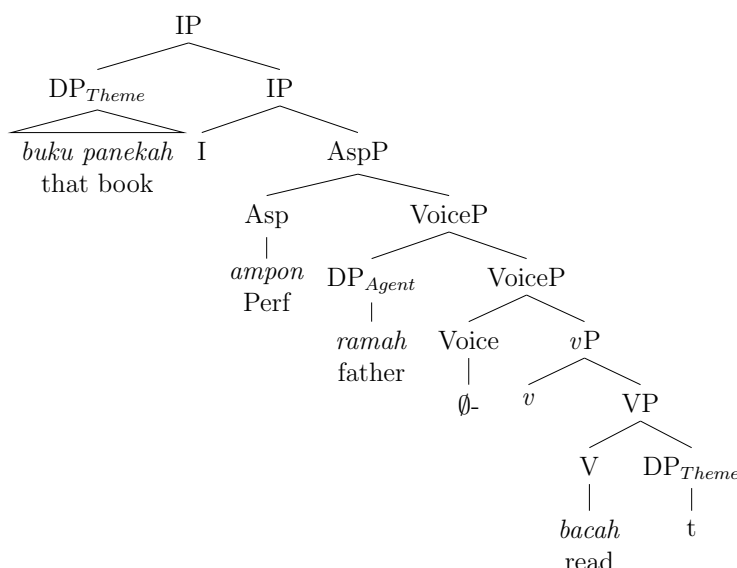
⁵ Under some circumstances, null prepositions are possible in Madurese, but in all cases where the preposition may be null, an overt preposition is also licit.

(13) Polite OV

Buku panekah (ampon) ramah (*ampon) bacah.
 buku panekah (ampon) ramah (*ampon) Ø-bacah.
 book that.Pol Perf.Pol father.Pol Perf.Pol OV-call.Pol-Appl
 ‘Father (already) read that book.’

The aspectual morpheme *ampon* cannot occur to the right of the OV Agent, whereas in active voice it appears between the Agent and the verb. The position of the OV Agent is that of external arguments, which I assume to be Spec,VoiceP (following Legate 2014):

(14) Object Voice Structure



Unlike AV clauses in which the external argument raises to the position of surface subjects, the OV Agent remains in its base-generated thematic position while the Patient/Theme is promoted to subject. Thus both Agent and Patient/Theme retain prominence as core arguments.

The inability of any element to intervene between Agent and verb follows from the position of the Agent, which remains in Spec,VoiceP. In languages like Madurese, where the OV Agent occurs preverbally, this means that modals and aspectual morphemes always precede the Agent. Additionally, an interesting morphophonological effect, akin to affixation, also occurs in object voice. My consultants report that a speech pause is unnatural, or even impossible, between the OV Agent and the following verb. Placing contrastive stress or emphasis on the Agent is also unnatural; emphatic stress occurs instead on the verb. The Agent and the verb thus appear to form a single phonological unit in object voice.

The OV Agent is said to be immobile (Legate 2014) since it does not move from its thematic position; WH-movement is impossible for an OV Agent. In

object voice, questions (15a) and relative constructions (15b), which both employ a pseudo-cleft construction, are ill-formed (traces indicate the thematic position from which the Agent has been moved):

(15) Polite OV

- a. *Paserah se potraepon ampon t tembhalih?
 paserah se potra-epon ampon Ø-tembhal-ih
 who.Pol Relson.Pol-Def.Pol Perf.Pol OV-call.Pol-Appl
 ‘Who called his son?’
- b. *Ramah se potraepon ampon t tembhalih
 ramah se potra-epon ampon Ø-tembhal-ih
 father.Pol Relson.Pol-Def.Pol Perf.Pol OV-call.Pol-Appl
 ‘The father who called his son’
- c. *Potraepon ampon paserah tembhalih?
 potra-epon ampon paserah Ø-tembhal-ih?
 son.Pol-Def.Pol Perf.Pol who.Pol OV-call.Pol-Appl
 ‘Who called his son?’

In general, the preferred method for forming questions in Madurese is WH-in-situ. In OV however, in-situ questioning of the Agent is disallowed (15c). The inability to form WH-in-situ questions, combined with the inability of the Agent to undergo WH-movement, has the effect that in Madurese object voice, Agent questions are not possible at all.

In some languages with an object voice,⁶ the OV Agent is restricted to a specific subset of DPs. In Madurese, the OV Agent is restricted to personal polite pronouns and some kinship terms such as *ramah* ‘father,’ *ebhu* ‘mother,’ *ale* ‘younger sibling.’ Other DPs are disallowed as OV Agents.⁷ Compare the availability of full DPs such as ‘Mr. Tono’ and ‘the teacher’ in active voice (16) with the ungrammaticality of the equivalent object voice clauses (17):

(16) Polite AV

- a. Pak Tono nembhalih potraepon.
 Pak Tono N-(t)embhal-ih potra-epon
 Mr Tono AV-call.Pol-Appl son.Pol-Def.Pol
 ‘Mr. Tono called his son.’
- b. Ghuruh panekah macah buku.
 ghuruh panekah m-(b)acah buku.
 teacher that.Pol AV-read book
 ‘The teacher read a book.’

⁶ But not all, e.g. Acehnese (Legate 2014) and colloquial Jakarta Indonesian (Sneddon 2006).

⁷ There is some interspeaker variation with regard to the acceptability of certain DPs as OV Agent. It is clear that not all DPs are acceptable. Polite pronouns are generally accepted by all speakers, but the acceptability of kinship terms and names may vary.

(17) Polite OV

- a. *Potraepon ampon Pak Tono tembhalih.
potra-epon ampon Pak Tono Ø-tembhal-ih.
son-Def.Pol Perf.Pol Mr Tono OV-call-Appl
- b. *Buku panekah ampon ghuruh panekah bacah.
buku panekah ampon ghuruh panekah Ø-bacah
book that.Pol Perf.Pol teacher that.Pol OV-read

Turning now to the familiar register of Madurese, I find that the object voice is not available. The expected object voice word order and null voice marking are given in the familiar register below:

(18) Familiar

- a. *Ana'eng la engko' katoeh.
Ana'-eng la engko' Ø-kato-ih
child-Def Perf 1s.Fam OV-call-Appl
'I called his child.'
- b. *Buku jiyah ale' bacah.
buku jiyah ale' Ø-bacah.
book that yngr.sibling OV-read
'Younger brother read that book.'

Both the word order, as well as the bare verb, are judged ungrammatical in the sentences in (18). Other combinations of pronominal Agents and bare verbs are also unacceptable.

I conclude that in the familiar register, only active and passive voices are available. Thus the grammars of the polite and familiar registers of Madurese are differentiated by the number of available voices: the polite register has a three-voice system while the familiar register has a two-voice system. I attribute this difference to the inventory of available Voice⁰ heads in each register (see (14)). The Voice⁰ head that determines object voice, and which is realized as Ø-, is available in polite speech, but not in familiar speech.

3. Register and DP Extraction

In this section I show that the grammatical differences between the registers extends to the pattern of DP extraction as well. I first discuss the ability of subject DPs and object DPs to be extracted from the clause in the polite register. This pattern is then contrasted with the extraction facts in the familiar register.

3.1. Polite Madurese: Subject and Object Extraction

In Madurese, DP extraction is possible in a pseudo-cleft or relative employing the relative morpheme *se* (Davies 2010). First, the grammatical subject of an active clause may be freely extracted (traces show the position from which a surface subject has been moved):

(19) Polite AV

- a. Ramah se t ampon macah buku panekah.
ramah se ampon N-(b)acah buku panekah.
father.Pol Rel Perf.Pol AV-read book that.Pol
'It was Father who read that book.'
- b. Ramah se t aparengaghih buku panekah dha' kancahepon.
ramah se a-pareng-aghih buku panekah dha' kancah-epon.
Father.Pol Rel AV-give.Pol-Applbook that.Pol to friend-Def.Pol
'It was Father who gave the book to his friend.'

The surface subject may likewise be questioned via a clefted WH word:

(20) Polite AV

- Paserah se t ampon macah buku panekah?
paserah se ampon N-(b)acah buku panekah?
who.Pol Rel Perf.Pol AV-read book that.Pol
'Who read that book?'

The ability to extract the subject also holds for passive and object voices. If object shift has first occurred, the DP that is the surface subject may be extracted:

(21) Polite OV

- Buku panekah se t ampon ramah bacah.
buku panekah se ampon ramah Ø-bacah.
buku that.Pol Rel Perf.Pol father.Pol OV-read
'It was that book which father read.'

(22) Polite PV

- Ponapah se t ebacah?
ponapah se e-bacah?
what.Pol Rel PV-read
'What was read?'

Note that in all cases of subject extraction, no change occurs on the verb; verbal morphology reflects the voice of the clause, and is required.

Next, extraction of a DP from object position, while retaining the voice morphology on the verb, results in ungrammaticality:

(23) Polite AV

- a. *Buku panekah se ramah ampon macah t .⁸
buku panekah se ramah ampon N-(b)acah
book that.Pol Rel father.Pol Perf.Pol AV-read
'It was that book which Father read.'
- b. *Buku panekah se ramah aparengaghih t dha' kancachepon.
buku panekah se ramah a-pareng-aghih dha' kancach-epon
book that.Pol Rel father.Pol AV-give.Pol-Appl to friend-Def.Pol
'It was that book which Father gave to his friend.'

Clefted object questions are likewise ungrammatical with voice morphology:

(24) Polite AV

- *Ponapah se ramah ampon macah t ?
ponapah se ramah ampon N-(b)acah
what.Pol Rel father.Pol Perf.Pol AV-read
'What did father read?'

However, if the same clauses employ a verb with null voice morphology, the object extraction is licit. The ungrammaticality of (23) and (24) contrasts with the licit examples in (25) and (26), which differ only in verbal morphology:⁹

(25) Polite AV

- a. Buku panekah se ramah ampon bacah t .
buku panekah se ramah ampon bacah
book that.Pol Rel father Perf.Pol read
'It was that book which Father read.'
- b. Buku panekah se ramah ampon parengaghih t dha' kancachepon.
buku panekah se ramah ampon pareng-aghih dha' kancach-epon
book that.Pol Rel father Perf. Pol give.Pol-Appl to friend-Def.Pol
'It is that book which Father gave to his friend.'

⁸ Examples (21) and (23) appear similar, but the relative order of the Agent and aspectual morphemes indicates whether the clause is in AV or OV (see Cole et al 2008). AV word order is Agent, aspect, verb; OV word order is aspect, Agent, verb.

⁹ The verbs in (25) and (26) are not shown with Ø- affixation, and the gloss does not indicate OV for the null affix. This is in order to differentiate the bare verb in AV extraction from the bare verb of OV; in the next section, however, the two phenomena are correlated.

- (26) Polite AV
Ponapah se ramah ampon bacah t ?
Ponapah se ramah ampon bacah
what.Pol Rel father.Pol Perf.Pol read
'What did father read?'

Thus, polite Madurese allows both subject DPs and object DPs to extract from the clause, but object extraction is only licit when the verb does not bear the active prefix. This extraction pattern is also attested in Indonesian (Cole and Hermon 1998; Cole et al 2008; Saddy 1991; Sneddon et al 2012; Soh 1998; Voskuil 1996) as well as several varieties of Jambi Malay (Yanti 2010). In these languages, extraction of object DPs also requires a bare verb. Furthermore, the extraction pattern is limited to DPs across these languages; adjuncts that originate low (post-verbal position) are not restricted from movement to a higher position in the clause.

3.2. Familiar Madurese: Subject Extraction Only

Turning to the familiar register, the extraction of subject DPs is possible, and like the polite register, voice morphology is retained on the verb:

- (27) Familiar AV
a. Daud se t la macah buku jiyah.
Daud se la N-(b)acah buku jiyah.
Daud Rel Perf AV-read book that
'It was David who read that book.'
b. Sapah se t ngatoeh ana'eng.
sawah se N-(k)ato-ih ana'-eng.
who Rel AV-call-Appl child-Def
'Who called his child?'

- (28) Familiar PV
Apah se t ebacah bi' Daud?
apah se e-bacah bi' Daud?
what Rel PV-read by Daud
'What was read by David?'

However, object extraction is not possible in the familiar register. Objects cannot be extracted with the active prefix (29a), nor can be extracted with a bare verb (29b):

- (29) Familiar AV
a. *Buku jiyah se Daud la macah t .
buku jiyah se Daud la N-(b)acah
book that Rel Daud Perf AV-read

- b. *Buku jiyah se Daud la bacah t.
 buku jiyah se Daud la bacah
 book that Rel Daud Perf read
 ‘It is that book which David read.’

Therefore, familiar Madurese allows DPs to extract only from subject position, and object extraction is not possible at all. This pattern of subject-only extraction is common in other Austronesian languages as well.

It is worth noting that *bacah* ‘read’ is a verb that can be used in both polite and familiar registers, and appears throughout the examples above. The ability to extract the object cannot be determined by the verb root, as *bacah* allows object extraction in polite speech but not in familiar speech.

In summary, the extraction patterns described in this section demonstrate that Madurese exhibits two distinct extraction patterns within one language, according to register. That these differences have not been noted in previous descriptions of Madurese is likely due to the fact that the polite register has not been well documented. Previous grammars have largely utilized data from the familiar register.

The voice and extraction differences are summarized in the table below:

(30) Register differences in Madurese

	AV	PV	OV	Subject extraction	Object extraction (with bare verb)
Polite register	Yes	Yes	Yes	Yes	Yes
Familiar register	Yes	Yes	No	Yes	No

4. Discussion

In this section I discuss some of the implications resulting from the Madurese data presented in this paper, and consider an analysis that can account for the difference in DP extraction facts between the registers.

4.1. A Link Between Voice Morphology and DP Extraction

The Madurese AV prefix *N-* is cognate with an active nasal prefix in many languages of the area, including Indonesian. The AV prefix *a-*, however, appears historically unrelated to the nasal prefix. I take both *N-* and *a-* to be morphological instantiations of the same AV Voice⁰ head.¹⁰ As shown in examples throughout this paper, *a-* exhibits the same pattern as *N-* with regard to DP extraction. Both *N-* and

¹⁰ Davies (2010) reports that for some verbs, both AV prefixes are acceptable.

a- are required in basic active clauses; and both are incompatible with object DP extraction.

As previously mentioned, in basic AV sentences in Madurese, the active voice prefix is required on verbs. This is in contrast to languages like Indonesian, in which active prefixes are sometimes optional (Sneddon et al 2012); bare verbs sometimes occur in basic active clauses in Indonesian (e.g. when no DP extraction has occurred). But in Madurese, the verb in an active clause always has an active voice marking, *except* in the case that DP extraction has occurred from object position. In other words, active word order may only occur with a bare verb when the object has extracted, and only in Polite speech. Familiar clauses, which may not have bare verbs (because of the lack of object voice), also disallow object extraction. These facts suggest a direct relationship between voice, voice morphology and DP extraction.

4.2. Null Voice Marking and Object Extraction

The question that arises for Madurese is, what allows object DPs to extract in polite speech, while preventing the same extraction in familiar speech?

The availability of object voice appears to correlate with the ability to extract objects from an active clause: familiar Madurese does not have an object voice, and does not allow object extraction, while polite Madurese has an object voice, and allows object extraction with the null prefix. Several proposals have suggested a “blocking and deletion” analysis for this object extraction pattern (Cole and Hermon 1998, Saddy 1991, Voskuil 1996): the nasal AV prefix blocks the movement of an object over the verb, so the voice morpheme must be deleted for extraction to be licit. If such an analysis were extended to Madurese, deletion of the AV prefix would be licit in polite speech, while prefix deletion would be illicit in familiar speech; but the question remains why this difference should exist between the registers.

I adopt an analysis proposed for Indonesian in Cole et al 2008, which is described as voice agreement. This analysis assumes a phase-based syntax and includes two language-specific, morphological requirements for a Philippine-type voice system (which includes AV and OV, but not the canonical passive voice). First, the voice marker must reflect the position from which a DP has moved. AV morphology indicates that the Agent DP has been extracted from its thematic position¹¹ to subject position, while OV morphology indicates that an object DP has been extracted from vP. (Recall that in object voice, the VP-internal object has undergone object shift and occurs as surface subject, while the Agent remains in its thematic position.) Second, a morphological constraint prevents DPs from extracting when the morphosyntactic properties of voice conflict with the DP moving through the edge of the phase. Under this analysis, when an object is

¹¹ Cole et al (2008) take this to be Spec,vP while I have assumed Spec,VoiceP following Legate’s (2014) structure.

extracted from an AV clause, two DPs (Agent and object) must move through the phase edge, and these have conflicting properties: nominative/Agent and accusative/Patient, respectively. Thus object extraction is blocked in active clauses.

Cole et al (2008) show that in some varieties of Indonesian/Malay, all object extraction is indeed illicit. For these languages, both the nasal prefix and null prefix block should object movement. In the Malay variety of Sarang Lan however, voice and extraction patterns resemble polite Madurese, allowing object extraction with a bare verb; the null prefix does not exclusively indicate OV, but also indicates object extraction from an AV clause. In this case, the AV prefix blocks object extraction, but the null prefix is “defective” and does not block extraction.

Extending this analysis to the polite and familiar registers of Madurese, I propose that the two morphological requirements also obtain in Madurese. The null prefix in polite Madurese is akin to the defective null prefix in Sarang Lan Malay, and does not block object extraction. The AV prefix *N-* cannot occur with object extraction because it is incompatible with the conflicting properties of the extracted Agent and object. Thus in polite Madurese, objects may only extract if the verb bears a null prefix. In contrast, since familiar Madurese does not have object voice, null voice marking (bare verb) is consequently not possible. This prevents object extraction in the familiar register, since the defective null prefix which would allow extraction is not available.

5. Conclusion

I have demonstrated that Madurese has three morphosyntactically defined voices: an active voice, a canonical passive voice and an object voice. All three voices are available in the polite register, but the familiar register does not have object voice, resulting in only two voices. The grammatical difference between the registers also extends to the DP extraction patterns. The polite register patterns with languages that allow both subjects and objects to extract, where object extraction is only licit with null voice morphology. The familiar register patterns with languages that allow only subjects to extract. Finally, by adopting the voice agreement analysis in Cole et al 2008, the impossibility of object extraction in the familiar register is attributed to the unavailability of object voice and the null voice morphology which allows object movement.

References

- Arka, I Wayan. 2003. *Balinese morphosyntax: A lexical-functional approach*. Canberra: The Australian National University.
- Arka, I Wayan and Christopher D. Manning. 2008. Voice and grammatical relations in Indonesian: A new perspective. *Voice and Grammatical Relations in Austronesian Languages*, eds. Peter K. Austin and Simon Musgrave. Stanford: CSLI Publications.

- Chung, Sandra. 1976. On the subject of two passives in Indonesian. *Subject and topic*, ed. C.N. Li. New York: Academic Press.
- Cole, Peter and Gabriella Hermon. 1998. The typology of WH-movement: WH-questions in Malay. *Syntax* 1:221-258.
- Cole, Peter and Gabriella Hermon. 2008. Malay/Indonesian syntax from an Austronesian perspective: An introduction. *Lingua* 118:1431-1439.
- Cole, Peter, Gabriella Hermon and Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118:1500-1553.
- Dardjowidjojo, Soenjono. 1978. *Sentence patterns of Indonesian*. Honolulu: University of Hawaii Press.
- Davies, William D. 2010. *A grammar of Madurese*. Berlin/New York: De Gruyter Mouton.
- Davies, William D. 1999. Madurese and Javanese as strict word-order languages. *Oceanic Linguistics* 38:152-167.
- Guilfoyle, Eithne, Henrietta Hung and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10:375-414.
- Haspelmath, Martin, ed. 2001. *Language typology and language universals*. Berlin: Walter de Gruyter.
- Killiaan, H. N. 1897. *Madoereesche spraakkunst: Eerste stuk, inleiding en klankleer*. Batavia: Landsdrukkerij.
- Kulikov, Leonid. 2010. Voice typology. *Oxford Handbook of Linguistic Typology*, ed. Jae Jung Song. Oxford University Press.
- Legate, Julie Anne. 2012. Subjects in Acehnese and the nature of the passive. *Language* 88:495-525.
- Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. Cambridge, MA: MIT Press.
- Pawitra, Adrian. 2008. *Kamus lengkap bahasa Madura - Indonesia*. Jakarta: Dian Rakyat.
- Saddy, Douglas. 1991. WH scope mechanisms in Bahasa Indonesia. *MIT Working Papers in Linguistics* 15:183-218.
- Safioedin S. H., Asis, ed. 1977. *Kamus bahasa Madura - Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Sneddon, James. 2006. *Colloquial Jakartan Indonesian*. Canberra: Australian National University.
- Sneddon, James, K. Alexander Adelaar, Dwi N. Djenar and Michael C. Ewing. 2012. *Indonesian: A comprehensive grammar, 2nd ed.* New York: Routledge.
- Soh, Hooi Ling. 1998. Certain restrictions on A-bar movement in Malay. *Proceedings of the Third and Fourth Meetings of Austronesian Linguistics Association*, 295-308.
- Stevens, Alan. 1968. *Madurese phonology and morphology*. New Haven: American Oriental Society.

The Proceedings of AFLA 22

Voskuil, Jan. 1996. *Comparative morphology: Verb taxonomy in Indonesian, Tagalog and Dutch*. The Hague: Holland Academic Graphics.

Yanti. 2010. A reference grammar of Jambi Malay. Ph.D dissertation, University of Delaware.

FLOATING UNIVERSAL QUANTIFIER AS A BASE-GENERATED HEAD IN THE VP PERIPHERY*

Dong-yi Lin
Ghent University
lindongyi@gmail.com

The universal quantifier *niz* in Kavalan can immediately precede its DP associate or appear in a quantifier-floating construction where it is separated from its DP associate. This paper argues that floating *niz* is not derived from its non-floating counterpart as a result of stranding. They differ morphosyntactically and semantically. First of all, while non-floating *niz* in a negative sentence exhibits scope reconstruction, floating *niz* in a negative sentence induces scope-freezing effect. Secondly, floating *niz* should be analyzed as a full-fledged verb, whereas non-floating *niz* is a nominal modifier. Thirdly, floating *niz* is not sensitive to A/A' distinction. Fourthly, floating and non-floating *niz* can co-occur in a sentence. Finally, floating *niz* can receive interpretations that are not available to non-floating *niz*. The differences between floating and non-floating *niz* indicate that they are derivationally distinct. The fact that floating *niz* is morphosyntactically realized as a verb further suggests that it is base-generated in a functional head in the VP periphery.

1. Introduction

Like English *all* in (1) and French *tous* in (2), the universal quantifier *niz* 'all' in Kavalan, an Austronesian language in eastern Taiwan, can be immediately adjacent to its DP associate (3a, b) or appear in a quantifier-floating construction where it is separated from its DP associate (3c).¹

- (1) a. [All the teachers] have finished grading finals.
b. [The teachers] have all finished grading finals.

- (2) a. [Tous les enfants] ont vu ce film.
all the children have seen this movie
b. [Les enfants] ont tous vu ce film.
the children have all seen this movie

* This research is funded by Fonds Wetenschappelijk Onderzoek (FWO: 2009-Odyseus-Haegeman-G091409). I'd like to thank the audience at AFLA 22 for their comments and suggestions. I am grateful to Abas and Ngengi for teaching me their language.

¹ Glossing conventions in this paper follow the Leipzig Glossing Rules. Additional glossing conventions are as follows: AT: Actor focus; AT: Actor topic; AV: Agent voice; DM: Discourse marker; LNK: Linker; NCM: Non-common noun marker; PV: Patient voice; TT: Theme topic.

- (3) a. m-lizaq tu wasu [ya m-niz sunis].
 AV-like OBL dog ABS AV-all child
 ‘All the children like dogs.’²
- b. m-lizaq tu wasu [ya sunis m-niz].
 AV-like OBL dog ABS child AV-all
 ‘All the children like dogs.’
- c. m-niz m-lizaq tu wasu [ya sunis].
 AV-all AV-like OBL dog ABS child
 ‘The children all like dogs.’

There are two major syntactic analyses of quantifier floating: Q(quantifier)-stranding and base-generation. On the Q-stranding approach proposed by Sportiche (1988), quantifier floating results from the stranding of a universal quantifier in an intermediate position where its DP associate passes (Bošković 2004; Giusti 1990; Shlonsky 1991). As depicted in (4), a universal quantifier and its DP associate originally form a QP constituent but the DP can undergo movement on its own. This results in the separation of the quantifier from its DP associate on the surface.

In contrast, advocates of the base-generation approach argue that a floating quantifier should be analyzed as a base-generated adverbial in the left periphery of VP (Baltin 1995; Doetjes 1997; Torrego 1996; Williams 1982). On this analysis, the QP headed by a floating quantifier is adjoined to VP or other functional projections in the fine-grained IP cartography.

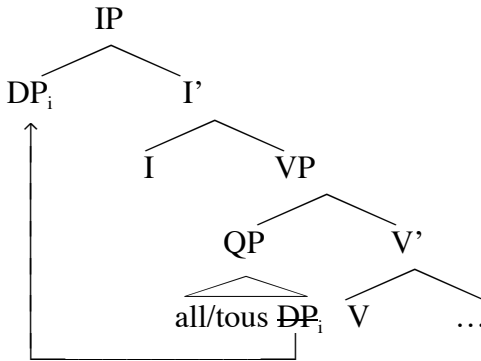
² The non-floating universal quantifier and its adjacent DP in (3a) and (3b) form a constituent. For example, they can be relativized together, as shown in (ib).

- (i) a. m-liyam tu sudad [ya m-niz pataqsian].
 AV-read OBL book ABS AV-all student
 ‘All the students read books.’
- b. [m-niz pataqsian] ya m-liyam=ay tu sudad.
 AV-all student ABS AV-read=REL OBL book
 ‘Those who read books are all the children.’

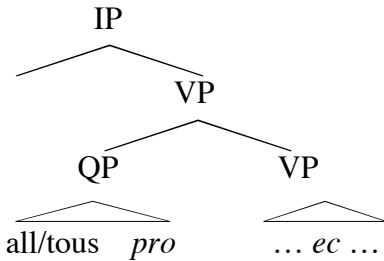
In contrast, the universal quantifier in Malagasy and its DP associate (*daholo* + DP) do not form a constituent despite their linear adjacency. As illustrated by the ungrammaticality of (iib), they cannot be topicalized together.

- (ii) Malagasy (Potsdam 2009: 769, 770)
- a. [[namaky ilay boky] daholo] ny mpianatra.
 read.AT that book all the students
 ‘The students all read that book.’
- b. *daholo ny mpianatra dia namaky ilay boky.
 all the student TOPIC read.AT that book
 (Intended for ‘All the students, they read that book.’)
- c. ny mpianatra dia namaky ilay boky daholo.
 the student TOPIC read.AT that book all
 ‘The students, they all read that book.’

(4) Q-stranding



(5) floating quantifier as a base-generated adverbial



The present paper aims to test the two approaches to quantifier floating against the distributional and morphosyntactic properties of Kavalan *niz*. It will be demonstrated that floating *niz* and non-floating *niz* exhibit semantic and morphosyntactic differences and are thus derivationally distinct. Quantifier floating of *niz* does not result from Q-stranding. Nevertheless, unlike English and French, where floating *all* and *tous* are an adverbial adjunct, floating *niz* is base-generated in a functional head position in the VP periphery.

Section 2 will first show that floating *niz* and non-floating *niz* observe the same restrictions regarding the type of DP/NP they can quantify over. Their similarities can be easily explained by the Q-stranding analysis of floating *niz*. However, section 3 will argue that the merits of the Q-stranding analysis are outweighed by the morphosyntactic and semantic differences between floating *niz* and non-floating *niz*. Another piece of evidence against the Q-stranding analysis concerns the word order derivation of a Q-floating construction. Section 4 concludes the study.

2. Floating *Niz* as a Stranded Q

On the Q-stranding analysis, floating *niz* is derived from non-floating *niz* and thus the two are expected to observe the same restriction on the type of DP/NP they can quantify over. Specifically, a DP that cannot co-occur with non-floating *niz* should not be able to bind floating *niz* either. This analysis also predicts that a DP that

cannot undergo movement should not be able to bind floating *niz*. These two predictions are borne out.

2.1. Personal Proper Names and *Niz*

The Q-stranding analysis predicts that floating *niz* and non-floating *niz* should observe the same restriction on the type of DP/NP they can quantify over. As indicated by the contrast between (6a) and (6b), personal proper names can co-occur with non-floating *niz* only when they are preceded by *qani-*, which denotes a group of people. Likewise, personal proper names can bind floating *niz* only when they are preceded by *qani-*, as illustrated in (7).

(6) a. *m-qila tu sunis [ya m-niz ti-buya, ti-imuy, ti-utay].
 AV-scold OBL child ABS AV-all NCM-Buya NCM-Imuy NCM-Utay

b. m-qila tu sunis [ya m-niz qani-buya, ti-imuy, ti-utay].
 AV-scold OBLchild ABS AV-all group-Buya NCM-Imuy NCM-Utay
 ‘Buya, Imuy, and Utay all scold children.’

(7) a. *m-niz m-qila tu sunis [ti-buya, ti-imuy, ti-utay].
 AV-all AV-scold OBL child NCM-Buya NCM-Imuy NCM-Utay

b. m-niz m-qila tu sunis [qani-buya, ti-imuy, ti-utay].
 AV-all AV-scold OBL child group-Buya NCM-Imuy NCM-Utay
 ‘Buya, Imuy, and Utay all scold children.’

The examples in (6) and (7) demonstrate that *qani-* must be attached to personal proper names when they are quantified over by either floating *niz* or non-floating *niz*. The fact that they share the same restriction can be attributed to the derivational history of floating *niz*, which originally forms a constituent with its DP associate and is later separated from it as a result of stranding.

2.2. DP-Movability and Quantifier Floating: Strong vs. Weak Quantifiers

The Q-stranding analysis also predicts that a DP that cannot undergo movement should not be able to bind a floating quantifier either. Before applying this test to Kavalan, I need to first discuss the distributional contrast between the strong quantifier, *niz* ‘all’, and the weak quantifier, *mwaza* or *mazmun* ‘many, much’.³

The examples in (8) and (9) illustrate the distributions of non-floating *niz* and *mwaza/mazmun* in different case positions. While *niz* can occur in an absolutive (8a) or ergative (8b) position, it cannot occur in an oblique position (8c). In contrast, *mwaza/mazmun* can only occur in an oblique position, as shown in (9).

³ *Mwaza* ‘many, much’ only modifies inanimate nouns, whereas *mazmun* ‘many, much’ is used when animate nouns are quantified over.

- (8) a. qa-qila-an-na ni imuy [ya m-niz sunis].
 QA-scold-PV-3ERG ERG Imuy ABS AV-all child
 ‘Imuy scolds all the children.’
- b. qa-qila-an-na [na patudan m-niz] ya sunis.
 QA-scold-PV-3ERG ERG teacher AV-all ABS child
 ‘All the teachers scold the child(ren).’
- c. *m-qila [tu m-niz sunis] ya ti-imuy.
 AV-scold OBL AV-all child ABS NCM-Imuy
 (Intended for ‘Imuy scolds all the children.’)
- (9) a. *t<m>anuz tu saku [ya mazmun wasu].
 <AV>chase OBL cat ABS AV.many dog
 (Intended for ‘Many dogs chase cats.’)
- b. *tanuz-an-na [na mazmun wasu] ya saku ’nay.
 chase-PV-3ERG ERG AV.many dog ABS cat that
 (Intended for ‘Many dogs chase that cat.’)⁴
- c. t<m>anuz [tu mazmun saku] ya wasu.
 <AV>chase OBL AV.many cat ABS dog
 ‘The dog chases many cats.’

The distributions of non-floating *niz* and *mwaza/mazmun* in different case positions are summarized in Table 1. The distributional contrast between the strong and weak quantifiers can be attributed to the default definiteness interpretation of a DP in the absolutive or oblique position. A Kavalan absolutive DP must be assigned a definite interpretation (Liao 2004) and is thus incompatible with a weak quantifier such as *mwaza/mazmun* ‘many, much’. In contrast, a Kavalan oblique DP receives an indefinite interpretation (Liao 2004), which conflicts with the semantics of a strong quantifier such as *niz* ‘all’.⁵

⁴ Kavalan *mazmun/mwaza* ‘many, much’ can appear in an absolutive or ergative position when it takes =*ay*.

- (i) a. t<m>anuz tu saku [ya mazmun=*ay* wasu].
 <AV>chase OBL cat ABS AV.many=*REL* dog
 ‘Many dogs chase cats.’
- b. tanuz-an-na [na mazmun=*ay* wasu] ya saku ’nay.
 chase-PV-3ERG ERG AV.many=*REL* dog ABS cat that
 ‘Many dogs chase that cat.’

⁵ The same distributional contrast between a strong quantifier and a weak quantifier has been observed in Seediq (Henningsson and Holmer 2008).

- (i) Seediq (Henningsson and Holmer 2008: 29)
- a. wada=*mu* puq-un kana bunga di.
 PST=1SG.ERG eat-PV all sweet.potato PRF
 ‘I ate all the sweet potatoes.’

Table 1. Strong and weak quantifiers in different case positions

	<i>m-niz</i> ‘all’	<i>mwaza/mazmun</i> ‘many, much’
ABS	✓	✗
ERG	✓	✗
OBL	✗	✓

As for floating *niz*, it must precede the lexical verb and be bound by the absolutive DP.⁶ As illustrated in (10), floating *niz* must quantifier over an absolutive DP, but not an oblique or ergative DP.

- (10)a. *m-niz m-liyam tu sudad [ya sunis].*
 AV-all AV-read OBL book ABS child
 ‘All the children read (a) book(s).’
 (NOT ‘The child(ren) read all the books.’)
- b. *m-niz qibasi-an-na na sunis [ya qudus].*
 AV-all wash-PV-3ERG ERG child ABS clothes
 ‘The child(ren) washed all the clothes.’
 (NOT ‘All the children washed the clothes.’)

-
- b. **mnekan=ku kana bunga di.*
 AV.PST.eat=1SG.NOM all sweet.potato PRF
 (Intended for ‘I ate all the sweet potatoes.’)

(ii) Seediq (Henningson and Holmer 2008: 27)

- a. **m-n-eyah hini egu preko.*
 AV-PST-come here much mosquito
 (Intended for ‘Many mosquitoes came here.’)
- b. *m-n-ari=ku egu blebul.*
 AV-PST-buy=1SG.NOM much banana
 ‘I bought many bananas.’

⁶ Malagasy *daholo* and Tsou *acuh* are also bound by a c-commanding DP, usually (though not always) the absolutive/nominative subject.

(i) Malagasy (Potsdam 2009: 769, 770)

- a. [[*namaky ilay boky*] *daholo*] *ny mpianatra.*
 read.AT that book all the students
 ‘The students all read that book.’
- b. *[[*hovakian’ ny mpianatra*] *daholo*] *ilay boky.*
 read.TT the students all that book
 (Intended for ‘The students all read that book.’)

(ii) Tsou (Chang 2002:331)

- mo acuh-u eobak-o ta ’o’oko ’e mamameoi.*
 AF all-AF beat-AF OBL children NOM old.men
 ‘These old men all beat the children.’

In contrast, *mwaza/mazmun* cannot appear in a Q-floating construction and be separated from its DP associate, as demonstrated by the contrast between (11a) and (11b).

- (11)a. m-Ramaz [tu mwaza tamun] ti-abas
 AV-cook OBL AV.many dish NCM-Abas
 ‘Abas cooks many dishes.’
- b. *mwaza m-Ramaz tu tamun ti-abas
 AV.many AV-cook OBL dish NCM-Abas

The distributional properties of *niz* and *mwaza/mazmun* seem to constitute empirical evidence for the Q-stranding analysis, which can relate their distributions to the well-known extraction restriction in Austronesian languages. On this analysis, there are two reasons why floating *niz* must be bound by an absolutive DP. First of all, as *niz* cannot co-occur with an oblique DP (8c), floating *niz* cannot have originated from an oblique position. Secondly, as only an absolutive DP can undergo extraction in Kavalan (Lin 2013), floating *niz* cannot have originated from an ergative position either. An ergative or oblique DP is unable to move to further induce stranding. The Q-stranding analysis also explains why *mwaza/mazmun* cannot float. As illustrated in (9), *mwaza/mazmun* can co-occur with an oblique DP, but not an absolutive DP. While an absolutive DP can move and is thus able to strand its modifying quantifier, as in the case of *niz*, an oblique DP cannot undergo movement, let alone strand *mwaza/mazmun*.

To summarize, the Q-stranding analysis can account for the same restriction shared by floating and non-floating *niz* regarding their co-occurrence with personal proper names. Moreover, it can explain why floating *niz* must be bound by an absolutive DP and why *mwaza/mazmun* cannot float. Nevertheless, as will be argued in section 3, the merits of the Q-stranding analysis are outweighed by the semantic and morphosyntactic differences between floating and non-floating *niz*. The fundamental differences between the two suggest that they are derivationally distinct.

3. Floating *Niz* as a Base-Generated Functional Head in the VP Periphery

Section 3.1 will demonstrate that floating *niz* is not equivalent to non-floating *niz*. They differ morphosyntactically and semantically and are thus derivationally distinct. Section 3.2 will further argue that the Q-stranding analysis fails to derive the correct word order of a Q-floating construction in Kavalan. The word order of floating *niz*, however, is not a problem for the base-generation analysis.

3.1. Semantic and Morphosyntactic Differences between Floating and Non-Floating *Niz*

3.1.1. Scope Interaction between *Niz* and Negation

According to Bošković (2004), English floating *all* induces scope-freezing effect in a negative sentence, whereas its non-floating counterpart exhibits scope reconstruction. As illustrated in (12a) and (12b), the scope of floating *all* in a negative sentence is unambiguous. When it follows the negation morpheme, negation must have a wider scope; when it precedes the negation morpheme, negation must have a narrower scope. In contrast, the scope of non-floating *all* in a negative sentence is ambiguous, as illustrated in (12c).

- (12)a. The students don't all speak Chinese. (not > all)
 b. The students all don't speak Chinese. (all > not)
 c. All the students don't speak Chinese. (all > not, not > all)

Likewise, non-floating *niz* in a negative sentence exhibits scope reconstruction. Either non-floating *niz* or the negation morpheme *mai* in (13a) can take wider scope over the other. Floating *niz* in a negative sentence, however, induces scope-freezing effect. (13b) and (13c) are both unambiguous. The linear order between floating *niz* and the negation morpheme *mai* determines their relative scope.

- (13)a. mai qibasi-an-na ni imuy [ya m-niz qudus].
 NEG wash-PV-3ERG ERG Imuy ABS AV-all clothes
 'Imuy didn't wash all the clothes.' (NEG > all; all > NEG)
 b. m-niz mai qibasi-an-na ni imuy [ya qudus].
 AV-all NEG wash-PV-3ERG ERG Imuy ABS clothes
 'Imuy didn't wash all the clothes.' (*NEG > all; all > NEG)
 c. mai m-niz qibasi-an-na ni imuy [ya qudus].
 NEG AV-all wash-PV-3ERG ERG Imuy ABS clothes
 'Imuy didn't wash all the clothes.' (NEG > all; *all > NEG)

3.1.2. Morphological Differences between Floating and Non-Floating *Niz*

Floating and non-floating quantifiers in some languages, e.g., Dutch and Korean, take different morphological forms, as illustrated in (14) and (15).

- (14) Dutch (Doetjes 1997: 210-211)
 a. [De kinderen] zijn allemaal gekomen.
 the children are all come
 'The children have all come.'
 b. *[Allemaal (de) kinderen] zijn gekomen.
 c. [Alle kinderen] zijn gekomen.
 all children are come
 'All of the children have come.'

(15) Korean (Jaehoon Choi, personal communication)

- a. Haksayng-tul-i [chayk-ul] modu ilk-ess-ta.
 student-PL-NOM book-ACC all read-PST-DECL
 ‘The students read all the books.’
- b. Haksayng-tul-i [modu-n chayk-ul] ilk-ess-ta.
 student-PL-NOM all-ADJ book-ACC read-PST-DECL
 ‘The students read all the books.’

Floating and non-floating *niz* also differ morphologically. While non-floating *niz* is a nominal modifier, floating *niz* should be analyzed as a full-fledged verb. First of all, floating *niz*, but not its non-floating counterpart, can take the imperative suffix, as shown by the contrast between (16a) and (16b). Secondly, floating *niz*, but not non-floating *niz*, can take the causative prefix, as illustrated in (16c). Finally, floating *niz* can be affixed with the patient voice marker (16d), whereas non-floating *niz* cannot (16e).

- (16)a. *niz-ika* *m-liyam* [ya *sudad*].
 AV-IMP.PV AV-read ABS book
 ‘Read all the books!’
- b. **m-liyam* [ya *niz-ika* *sudad*].
 AV-read ABS all-IMP.PV book
- c. *pa-niz=iku* [tu *sunis*] *pa-taqsi*.
 CAUS-all=1SG.ABS OBL child CAUS-study
 ‘I make all the children study.’⁷
- d. *niz-an-na=ti* *ni abas* q<m>an [ya *byabas*].
 all-PV-3ERG=PFV ERG Abas <AV>eat ABS guava
 ‘Abas ate all the guavas.’
- e. **qan-an-na=ti* *ni abas* [ya *niz-an byabas*].
 eat-PV-3ERG=PFV ERG Abas ABS all-PV guava

Moreover, the voice markers on floating *niz* are verb-defining *v*, which can determine the argument structure of a sentence. While PV *-an* by itself can assign an external argument and an affected theme, AV *m-* cannot (Lin 2013, 2015). PV-marked *niz* in (17a) is thus grammatical even without a lexical verb, but this is not true of AV-marked *niz* in (17b).

- (17)a. *niz-an-na=ti* *ni abas* [ya *byabas*].
 all-PV-3ERG=PFV ERG Abas ABS guava

⁷ In (16c), not only the universal quantifier but also the lexical verb is affixed with the causative marker *pa-*. The doubling of the causative marker is reminiscent of “prefix harmony” or “anticipating sequence” in Siraya (Adelaar 1997; Tsuchida 2000). Whether this Kavalan example can be subsumed under this phenomenon requires further investigation.

- ‘Abas ate/used up all the guavas.’
- | | | | | | |
|----|--------|-----|--------|-----|---------|
| b. | *m-niz | tu | byabas | [ya | sunis]. |
| | AV-all | OBL | guava | ABS | child |

All these facts suggest that floating *niz* should be analyzed as a full-fledged verb that can move to *v*. This empirical generalization is incompatible with the Q-stranding analysis, which predicts that the stranded quantifier should be embedded inside a specifier position and cannot undergo head movement to *v*. Instead, as low adverbials in the VP periphery are all realized as a verb in Kavalan (Chang 2006; Lin 2014), the facts presented in (16) and (17) corroborate the analysis of floating *niz* as a base-generated head in the VP periphery below *v*.⁸ As illustrated in (18), low adverbials in Kavalan share the same distributional and morphosyntactic properties as a verb.

- | | | | | |
|--------|-------------------------------------|----------------|-----|--------|
| (18)a. | paqanas-an-ku | m-liyam | ya | sudad. |
| | slowly-PV-1SG.ERG | AV-read | ABS | book |
| | ‘I read the book slowly.’ | | | |
| b. | m-duna | qan-an-ku | ya | ’may. |
| | AV-always | eat-PV-1SG.ERG | ABS | rice |
| | ‘I always eat (this kind of) rice.’ | | | |

3.1.3. Insensitivity to A/A’ Distinction

Sensitivity to A/A’ distinction has often been cited as evidence for the stranding analysis of quantifier floating. As indicated by the contrast between (19) and (20), a DP that undergoes A’-movement cannot strand a quantifier, unless it first undergoes short A-movement (Bobaljik 2003; Déprez 1989). A floating quantifier is only licensed in an A-chain due to its immediate adjacency to the DP-trace of its associate.

- (19) (Bobaljik 2003)
- The runners_i seem to themselves [_{t_i} to be moving very slowly].
 - The lions_i might all seem (to you) [_{t_i} to have large teeth].
 - The lions_i might all have been seen _{t_i} (by the tourists).
- (20) (Bobaljik 2003)
- *the professors who Taylor will have all met before the end of term
 - *These professors, Taylor will have all met before the end of term.
 - *Which professors will Taylor have all met before the end of term?

This argument, however, does not apply to Kavalan *niz*. In (21), the operator that floating *niz* quantifies over undergoes A’-movement without

⁸ cf. the high base-generation site of Malagasy *daholo* proposed by Koopman (2005)

incurring ungrammaticality. On the base-generation approach to floating *niz*, the grammaticality of (21) is expected, as floating *niz* and its DP associate never form a constituent and thus the movement type of the associate does not determine whether *niz* can float or not.

- (21)a. byabas [_{RC} niz-an-na=ay ni buya m-Rasa]
 guava all-PV-3ERG=REL ERG Buya AV-buy
 ‘the guavas that Buya all bought’
- b. saku ya [_{Headless RC} niz-an-na=ay na wasu t<m>anuz].
 cat ABS all-PV-3ERG=REL ERG dog <AV>chase
 ‘What the dog chases all are the cats.’
- c. niana ya [_{Headless RC} niz-an-na=ay na wasu t<m>anuz]?
 what ABS all-PV-3ERG=REL ERG dog <AV>chase
 ‘What does the dog chase all?’

3.1.4. Co-occurrence of Floating and Non-Floating *Niz*

If floating *niz* and non-floating *niz* are not derivationally related, as claimed by the base-generation approach, it is predicted that they should be able to co-occur in a sentence. As shown in (22), this prediction is borne out.

- (22) niz-an-na ni abas q<m>an [ya m-niz byabas].
 all-PV-3ERG ERG Abas <AV>eat ABS AV-all guava
 ‘Abas ate up all the guavas.’

The co-occurrence of floating Q and non-floating Q can also be observed in other languages, e.g., Chinese (23), Korean (24), and Cebuano (25).

- (23) Chinese
 [suoyou de xuesheng] dou xihuan zhe tang ke.
 all DE student all like this CLF class
 ‘All the students like this class.’
- (24) Korean (Jaehoon Choi, personal communication)
 [modu-n haksayng-tul-i] modu chayk-ul ilk-ess-ta.
 all-ADJ student-PL-NOM all book-ACC read-PST-DECL
 ‘All the students read the book.’
- (25) Cebuano
 pulos ganahan ani nga libro [ang tanan nga mga bata].
 all like this LNK book ABS all LNK PL child
 ‘All the children like this book.’

Floating Q in (22) to (25) is not a pronounced copy of the corresponding non-floating Q, as they are formally distinct from each other.

3.1.5. Interpretations of Floating *Niz*

Floating *niz* can receive interpretations that are not available to non-floating *niz*. Floating *niz* can be used as a quantifier/modifier of either entities/individuals or predicates. It can either mean ‘all’ or ‘completely; entirely’. For example, floating *niz* in (26a) can quantify over either the absolutive DP *ya qudus-ku* or the predicate *tengen*.

- (26)a. m-niz tengen ya qudus-ku.
 AV-all black ABS clothes-1SG.GEN
 ‘My clothes are completely black.’ or ‘All my clothes are black.’
- b. maqen-ika m-niz q<m>an ya byabas.
 indeed-IMP.PV AV-all <AV>eat ABS guava
 ‘Do eat up (all) the guavas.’ (‘consume entirely’)

Floating *niz* can also function as a quantifier of a set of propositions. It is similar to Chinese *dou*, which can quantify over a set of propositions denoted by an embedded question or a yes-no question (i.e., their possible answers), as illustrated in (27) (Cheng 1995; Cheng & Huang 1996). A corresponding Kavalan sentence is given in (28).

- (27) Chinese
- a. ni guyong shei, wo dou hui bang ta.
 you hire who I all will help him
 ‘No matter who you hire, I will help him.’
- b. ta nian bu nian dou gen wo mei guanxi.
 he study NEG study all with I NEG matter
 ‘It is none of my business whether or not he studies (this).’

- (28) t<m>ayta tu ti-tiana wasu zau nani, niz-an-na Raytunguz
 <AV>see OBL NCM-who dog this DM all-PV-3ERG bark
 ‘No matter who this dog sees, it barks (at him).’

3.1.6. Interim Summary

Section 3.1 has presented 5 empirical facts of *niz* that are opaque to the Q-stranding analysis: (1) scope-freezing effect (floating *niz*) vs. scope reconstruction (non-floating *niz*); (2) floating *niz* as a full-fledged verb; (3) the insensitivity of floating *niz* to A/A’ distinction; (4) co-occurrence of floating *niz* and non-floating *niz*; (5) floating *niz* as a quantifier of a set of propositions. The fundamental differences

between floating *niz* and non-floating *niz* suggest that they are derivationally distinct. Just like other low adverbials in Kavalan, floating *niz* is also morphosyntactically realized as a verb and should be analyzed as a base-generated functional head in the VP periphery.

3.2. Word Order of Floating *Niz*

This section will argue that the Q-stranding analysis also fails to derive the word order of a Q-floating construction in Kavalan.

Verb-initial order in Austronesian languages can be derived via either TP/VP fronting, e.g., Malagasy (Pearson 2005) and Seediq (Aldridge 2004, 2013), or verb raising, e.g., Tagalog (Aldridge 2004, 2012). On the TP/VP fronting approach, the absolutive subject moves to Spec, TopP to check the uninterpretable [D]-feature [*D] on Top. The verb-initial word order results from the subsequent movement of the remnant TP or predicate phrase to Spec, FocP.

The verb-initial order of Tagalog, however, is derived via verb raising (Aldridge 2004, 2012). In this language, the absolutive subject undergoes covert movement to the outer specifier of *v*P, where it receives a definite, presuppositional interpretation. There is no TP/VP fronting. Instead, the verb-initial order is derived via verb movement to T.

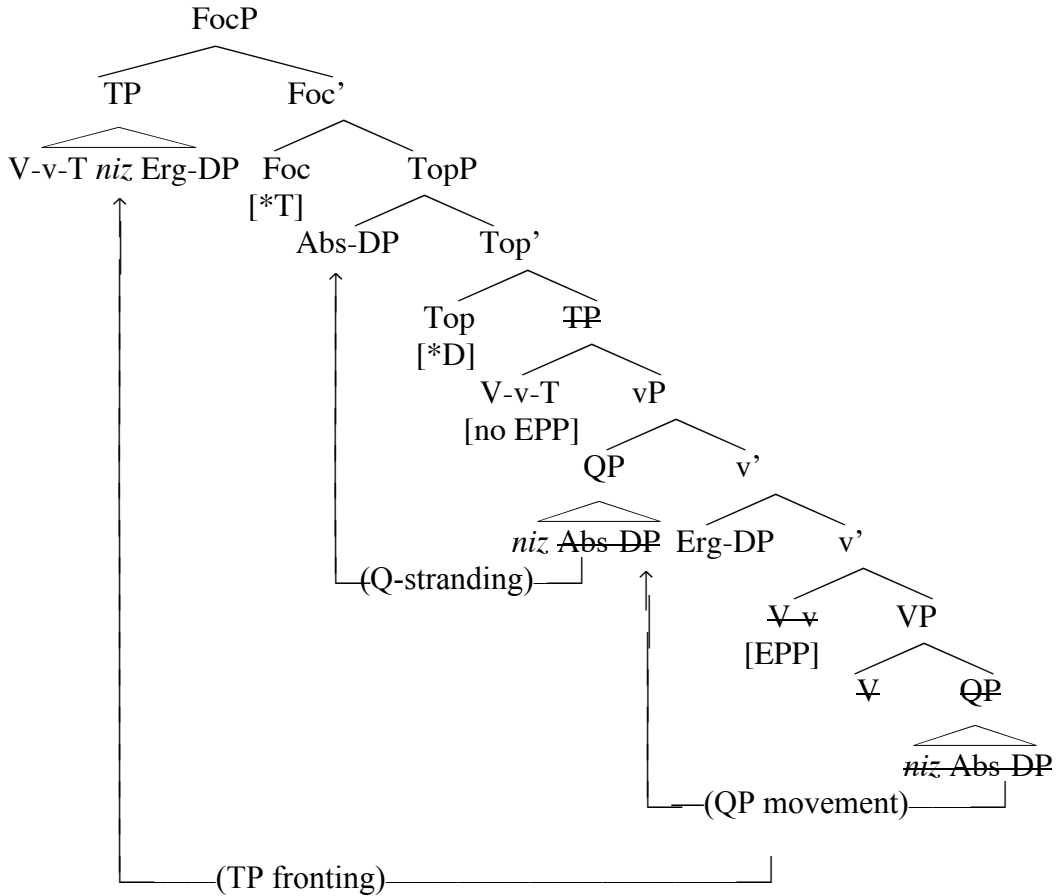
On the Q-stranding analysis, [*niz* Erg-DP V Abs-DP], as illustrated in (29), is not a possible word order in Kavalan regardless of how Kavalan verb-initial order is derived.

- (29) *niz-an-na=ti* *ni* *abas* *q<m>an* [*ya* *byabas*].
 all-PV-3ERG=PFV ERG Abas <AV>eat ABS guava
 ‘Abas ate all the guavas.’

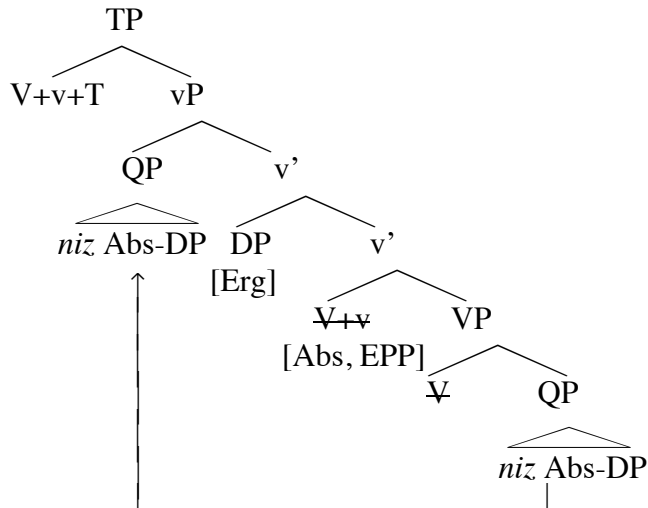
The derivation in (30) involves both quantifier stranding and TP fronting. The correct order of [*niz* Erg-DP V Abs-DP] cannot be derived. As shown in (31), on the analysis of Q-stranding, verb raising also fails to derive the word order of floating *niz*, which precedes the lexical verb in a Q-floating construction. If quantifier floating of *niz* in Kavalan results from quantifier stranding accompanied by DP movement, floating *niz* should never precede a verb, contrary to fact.

On the base-generation approach, floating *niz* is merged in a functional head position that is structurally higher than V. Its clause-initial position before V is expected no matter how Kavalan verb/predicate-initial order is derived. The derivations in (32) and (33) show that as a base-generated functional head between *v* and V, floating *niz* always precedes the lexical verb whether Kavalan verb-initial order is derived via TP fronting or verb raising.

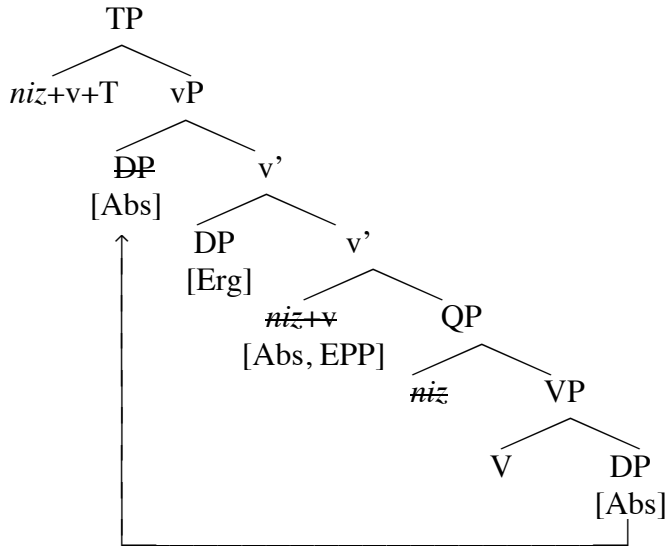
(30) TP fronting + Q-stranding → incorrect word order



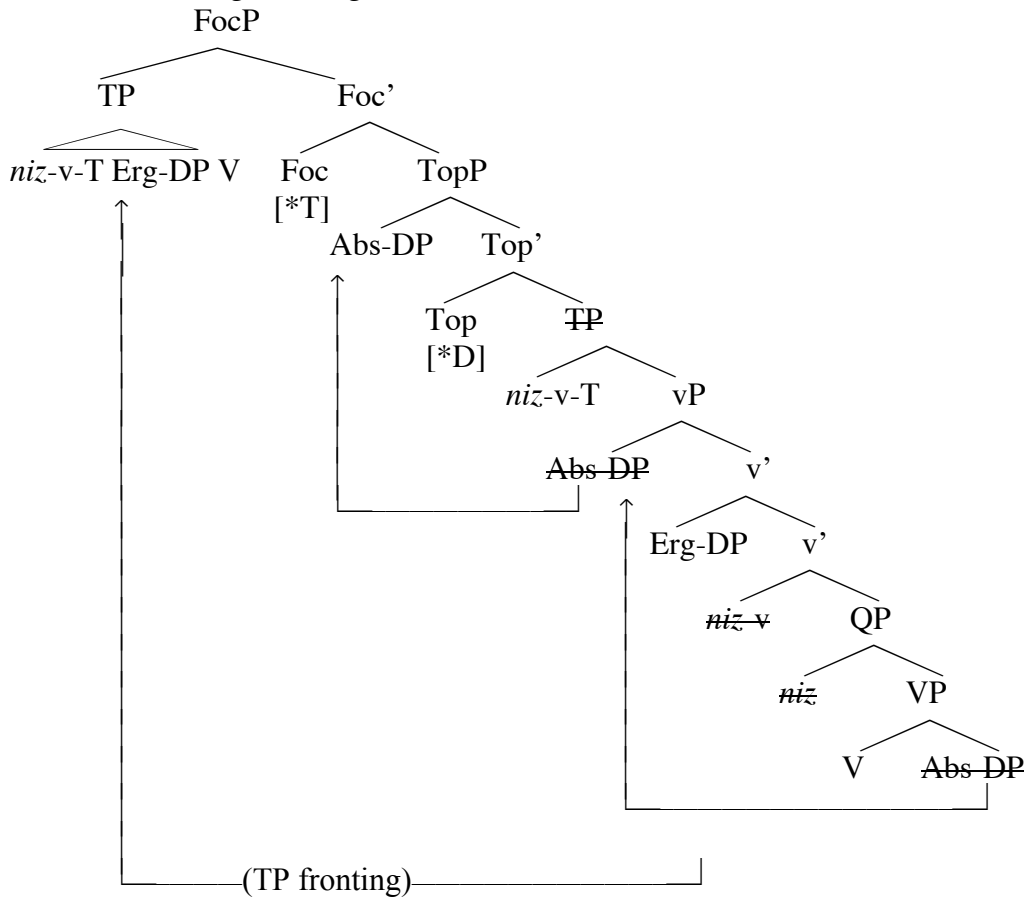
(31) verb raising + Q-stranding → incorrect word order



(32) verb raising + base-generated Q between *v* and V



(33) TP fronting + base-generated Q between *v* and V



4. Conclusion

Floating *niz* and non-floating *niz* in Kavalan differ both semantically and morphosyntactically. Floating *niz* is not derived from non-floating *niz* as a result of quantifier stranding. Non-floating *niz* is a nominal modifier, whereas floating *niz* is base-generated in a head position in the VP periphery below *v* and exhibits properties of a full-fledged verb just like other adverbial verbs in this language.

References

- Adelaar, K. Alexander. 1997. Grammar notes on Siraya, an extinct Formosan language. *Oceanic Linguistics* 36:362-397.
- Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral dissertation, Cornell University, Ithaca.
- Aldridge, Edith. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122:192-203.
- Aldridge, Edith. 2013. Wh-clefts and verb-initial word order in Austronesian languages. *Cleft structures*, eds. Katharina Hartmann and Tonjes Veenstra, 71-96. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Baltin, Mark R. 1995. Floating quantifiers, PRO, and predication. *Linguistic Inquiry* 26:199-248.
- Bobaljik, Jonathan David. 2003. Floating quantifiers: Handle with care. *The second Glot international state-of-the-article book: The latest in linguistics*, eds. Lisa Cheng and Rint Sybesma, 107-148. Berlin/New York: Mouton de Gruyter.
- Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language and Linguistic Theory* 22:681-742.
- Chang, Henry Yung-li. 2002. Distributivity, plurality, and reduplication in Tsou. *The Tsing Hua Journal of Chinese Studies, New Series* 32:327-355.
- Chang, Yung-li. 2006. The guest playing host: Adverbial modifiers as matrix verbs in Kavalan. *Clause Structure and Adjuncts in Austronesian Languages*, eds. Gärtner Hans-Martin, Paul Law, and Joachim Sabel, 43-82. Berlin: Mouton de Gruyter.
- Cheng, Lisa Lai-Shen. 1995. On *dou*-quantification. *Journal of East Asian Linguistics* 4:197-234.
- Cheng, Lisa Lai-Shen, and C.-T. James Huang. 1996. Two types of donkey sentences. *Natural Language Semantics* 4:121-163.
- Déprez, Viviane. 1989. On the typology of syntactic positions and the nature of chains: Move α to the specifier of functional projections. Doctoral dissertation, MIT, Cambridge.
- Doetjes, Jenny Sandra. 1997. Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English. Doctoral dissertation, Leiden University, The Hague.
- Giusti, Giuliana. 1990. Floating quantifiers, scrambling, and configurationality. *Linguistic Inquiry* 21.633-641.

- Henningsson, Lars-Åke, and Arthur Holmer. 2008. The distribution of quantifiers in Seediq. *Lund University Department of Linguistics and Phonetics Working Papers* 53:23-41.
- Liao, Hsiu-chuan. 2004. Transitivity and ergativity in Formosan and Philippine Languages. Doctoral dissertation. The University of Hawai'i, Manoa.
- Lin, Dong-yi. 2013. Interrogative constructions in Kavalan and Amis. Doctoral dissertation, University of Florida, Gainesville.
- Lin, Dong-yi. 2014. A cartographic approach to adverbials in Kavalan: Adverbial verbs and restructuring. Paper presented at 2014 Annual Meeting of the Linguistics Association of Great Britain. University of Oxford.
- Lin, Dong-yi. 2015. The syntactic derivations of interrogative verbs in Amis and Kavalan. *New advances in Formosan linguistics*, eds. Elizabeth Zeitoun, Stacy F. Teng, and Joy J. Wu, 253-289. Canberra: Asia-Pacific Linguistics.
- Koopman, Hilda. 2005. Malagasy imperatives. *Proceedings of AFLA XII*, 141-160. (*UCLA Working Papers in Linguistics* 12) University of California, Los Angeles.
- Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory* 23:381-457.
- Potsdam, Eric. 2009. Malagasy backward object control. *Language* 85:754-784.
- Shlonsky, Ur. 1991. Quantifiers as functional heads: A study of Quantifier Float in Hebrew. *Lingua* 84:159-180.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425-449.
- Torrego, Esther. 1996. On quantifier float in control clauses. *Linguistic Inquiry* 27:111-126.
- Tsuchida, Shigeru. 2000. Lexical prefixes and prefix harmony in Siraya. *Grammatical analysis: Morphology, syntax, and semantics – Studies in honor of Stanley Starosta*, eds. Videia P. De Guzman and Byron W. Bender, 109-128. Honolulu: University of Hawai'i Press.
- Williams, Edwin. 1982. The NP cycle. *Linguistic Inquiry* 13:277-295.

THE GRADUAL PATH TO VARIABLE REDUPLICATION: KAVALAN

Kuo-Chiao Lin
New York University
kcl299@nyu.edu

This paper presents a case study of Kavalan to show that variable bimoraic reduplication can be tackled in Serial Template Satisfaction (STS), a theory of reduplication couched in Harmonic Serialism, without stipulations of any weight conditions or multiple allomorphic templates. This paper also demonstrates that the proposed analysis can be extended to heavy syllable reduplication in Ilokano as well. The analysis shows that in STS the ultimate shape of a reduplicant in a given language is a result from the collective effort by the template and the general phonology of the language.

1. Introduction

This paper attempts to examine *Serial Template Satisfaction* (STS, McCarthy et al. 2012), a novel theory of reduplication situated in Harmonic Serialism (HS, McCarthy 2000, McCarthy 2010), in the context of Kavalan, a Formosan language which exhibits different bimoraic reduplicative shapes for the same continuative prefix, as shown in (1).

- (1) *Continuative reduplication in Kavalan*
- | | | | |
|----|-------------------------------------|-------------------|-----------------------|
| a. | m- <u>vi</u> .- <u>βa.u</u> .-βa.ut | ‘to keep fishing’ | CVV (Lin 2012: 1052) |
| b. | m- <u>u.ɿi</u> .-u.ɿiŋ | ‘to keep weeping’ | VCV (Lin 2012: 1052) |
| c. | <u>pu.ku</u> .-pu.kun | ‘to keep hitting’ | CVCV (Chang 2000: 60) |
| d. | <u>tum</u> .-tum.βəs | ‘to keep pulling’ | CVC (Lin 2012: 1057) |

Given that coda consonants in Kavalan are moraic (Lin 2012), we see in (1) that the reduplicant is bimoraic. If the stem’s first syllable is light, the reduplicant is (C)V(C)V, as in (1a-c). If, on the other hand, the stem’s first syllable is heavy, then the reduplicant is CVC, as in (1d). How STS addresses bimoraic reduplicative variants has not been thoroughly discussed in McCarthy et al. (2012). Since STS assumes that a reduplicative affix is an underlying template (Marantz 1982, McCarthy and Prince 1996 [1986]), three serious questions arise for the theory, as shown in (2).

- (2) a. How does STS account for languages like Kavalan, where a single reduplicative morpheme shows different shapes (with different stems)?

- b. What is the source of CVC reduplicants in the language? Is any weight requirement necessary? If so, should it be formulated as a heavy syllable template, e.g., $\sigma_{\mu\mu}$ (McCarthy and Prince 1996 [1986], see also Marantz 1982, Thurgood 1997), or as a constraint on the reduplicant, e.g., RED= $\mu\mu$ (McCarthy and Prince 1993b, see also Blevins 1996, Crowhurst 2004)?

This paper provides answers to these two questions by showing that STS can capture not only the reduplicative pattern in Kavalan, but also the emergence of what looks like a heavy syllable reduplicative template, solely by constraint interaction and the general phonology of the language. The central proposal is that there is no need to assume multiple allomorphic templates or weight requirements to account for the overall pattern. The paper is organized as follows. Section 2 gives a brief overview of STS, followed by demonstrations in Section 3 that constraint interaction alone is able to address the bimoraic reduplicative pattern in Kavalan. Section 4 extends the proposed analysis to heavy syllable reduplication in Ilokano, and Section 5 concludes the paper.

2. Serial Template Satisfaction

Serial Template Satisfaction (STS, McCarthy et al. 2012) extends Harmonic Serialism (McCarthy 2000, McCarthy 2010) to the realm of reduplication in addition to other phonological phenomena (McCarthy 2000, McCarthy 2008a, McCarthy 2008b, McCarthy 2010, Wolf 2008, Elfner 2009, Pruitt 2010, Kimper 2011, Lin 2014, inter alia). The theory assumes that reduplicative affixes are templates (Marantz 1982), which, below the prosodic word level, are syllables(σ) or feet(FT) (McCarthy and Prince 1996 [1986]).¹ In STS, a template X violates HEADEDNESS(X), as defined in (3).

- (3) HEADEDNESS(X)(HD(FT), HD(σ)) (Selkirk 1995)
Assign a violation mark for every constituent of type X that does not contain a constituent of type X-1 as its head.

The HD constraint can be fulfilled by applications of COPY(X) or INSERT(X) performed by GEN. The operation INSERT(X) inserts a prosodic constituent node of type X into the existing structure and incurs a violation of HD(X).² For example, given a foot template, INSERT(σ) inserts a syllable node to fill the foot template, resolving the violation of HD(FT), but incurs a violation of HD(σ), as in (4).

¹ The template can as well be a prosodic word, which may be fulfilled by COPY(FT) or COPY(*stem*) (McCarthy et al. 2012).

² As McCarthy et al. (2012: 180) point out, whether INSERT(X) violates faithfulness constraints is not important in STS. Note, however, any faithfulness violation as such entails a HD(X) violation, but not vice versa.

- (4) INSERT(X-1): $X = FT, X-1 = \sigma$
 FT-(ta.sa) \rightarrow (σ)-(ta.sa)
 *HD(FT) \checkmark HD(FT)
 *HD(σ)

COPY(X), on the other hand, is the sole source of reduplicative identity, eliminating correspondence constraints in BR-Correspondence (McCarthy and Prince 1995, McCarthy and Prince 1999). The COPY(X) operation copies a *string* of constituents of type X along with their content into the pre-existing prosodic structure. Since what is copied is a string, the constituents copied by a single application of COPY(X) are necessarily *contiguous* and of the type X. This is shown in (5).

- (5) COPY(X-1): $X = FT, X-1 = \sigma$
 FT-(ta.sa) \rightarrow (ta)-(ta.sa), (ta.sa)-(ta.sa), *(tas)-(ta.sa)

Since COPY(X) is defined to copy strings of elements of type X, a single application of COPY(X) incurs a violation of *COPY(X), *regardless of the number of the constituents contained in the copied string*. Thus, copying one syllable, e.g., FT-(ta.sa) \rightarrow (ta)-(ta.sa), and copying two syllables, e.g., FT-(ta.sa) \rightarrow (ta.sa)-(ta.sa), are equally unfaithful in terms of *COPY(σ). Moreover, the string-copying property of COPY(X) may produce a violation of COPY-LOCALLY(X), as defined in (6).

- (6) COPY-LOCALLY(X) (COPY-LOC, McCarthy et al. 2012:181)
 To a candidate produced by Copy(X), assign as many violations as there are Xs intervening between the original X string and its copy.

As illustrated in (7), the constraint replaces ANCHOR-BR (McCarthy and Prince 1995) and CONTIGUITY-BR (McCarthy and Prince 1999) in accounting for the “edge-in” effect (Marantz 1982) in reduplication.

- (7) a. *BR-Correspondence*

	ANCHOR-BR	CONTIGUITY-BR
$p_1i_2t_3b_4a_5r_6d_7u_8$		
i. $\rightarrow p_1i_2t_3b_4a_5-p_1i_2t_3b_4a_5r_6d_7u_8$		
ii. $p_1i_2t_3d_7u_8-p_1i_2t_3b_4a_5r_6d_7u_8$		3W($b_4a_5r_6$)
iii. $b_4a_5r_6d_7u_8-p_1i_2t_3b_4a_5r_6d_7u_8$	1W	

- b. *STS*

	COPY-LOC(<i>seg</i>)
$p_1i_2t_3b_4a_5r_6d_7u_8$	
i. $\rightarrow p_1i_2t_3b_4a_5-p_1i_2t_3b_4a_5r_6d_7u_8$	
ii. $b_4a_5r_6d_7u_8-p_1i_2t_3b_4a_5r_6d_7u_8$	3W($p_1i_2t_3$)

All in all, the tug-of-war among *COPY(X), HD(X), and other markedness constraints like FT-BIN (Prince 1980, McCarthy and Prince 1996 [1986]), for

example, determines not only whether phonological processes precede or follow applications of COPY(X), but also how the reduplicant is shaped.

3. Bimoraic Reduplication in Kavalan: An STS Analysis

The generalization for bimoraic reduplication in Kavalan, repeated below, is that if reduplication of the stem's first syllable is enough to meet the bimoraicity, then the said syllable is copied (e.g., CVC), as in (8a-c); otherwise, the reduplicant is (C)V(C)V, as in (8d).

(8) *Continuative reduplication in Kavalan*

- | | | | |
|----|-------------------|-------------------|-----------------------|
| a. | m-ki.-βa.u.-βa.ut | 'to keep fishing' | CVV (Lin 2012: 1052) |
| b. | m-u.ki.-u.kiŋ | 'to keep weeping' | VCV (Lin 2012: 1052) |
| c. | pu.ku.-pu.kun | 'to keep hitting' | CVCV (Chang 2000: 60) |
| d. | tum.-tum.βəs | 'to keep pulling' | CVC (Lin 2012: 1057) |

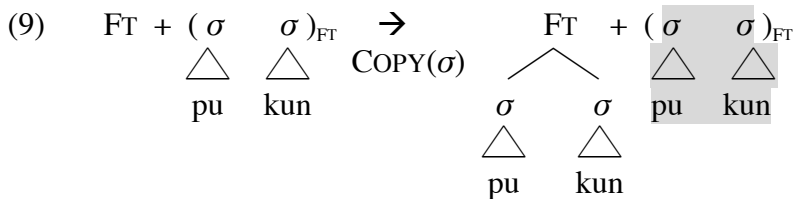
Consequently, a satisfactory STS analysis must explain the conditions responsible for the variable reduplicative shapes, closed syllable reduplication in particular. If a foot template will ultimately be realized as disyllabic, whereas a syllable template will surface as CV, or CVC if a certain heaviness requirement is specified (McCarthy et al. 2012: 224-225), then STS seems to require two moves to derive the reduplicative pattern in Kavalan. *First*, the theory might need to assume two allomorphic templates, a syllable and a foot, that are respectively responsible for monosyllabic and disyllabic reduplicative variants. *Second*, to derive the CVC variants, STS may have to impose some heaviness condition on the syllable template (cf., McCarthy et al. 2012: 197).

Two problems arise, however. *First*, given that the reduplicative variants in (8) all carry identical semantics, it is worth rethinking whether we really need to presume that some of the reduplicants start as a foot, while the others originate as a syllable. If the forms of the allomorphs are predictable from context, then lexically listing allomorphs would miss a generalization. *Second*, how the heaviness condition on the CVC variants can be formulated is unclear. Should the heaviness condition be stipulated on the template itself, namely $\sigma_{\mu\mu}$ (McCarthy and Prince 1996 [1986], see also Marantz 1982, Thurgood 1997), or should it be formulated as a constraint on the reduplicant, e.g., RED= $\mu\mu$ (McCarthy and Prince 1993b, see also Blevins 1996, Crowhurst 2004)?

I argue, instead, that neither multiple allomorphic templates nor weight conditions on the template are required for an STS analysis of the pattern in (8). To set the stage for the analysis, I need to determine whether the size of the template employed in the reduplicative process is a foot or a syllable. Since the reduplicant is bimoraic, the null hypothesis is that the template is a foot, which, I propose, may result in bimoraic variants, (C)V(C)V or CVC, due to FT-BIN. The question for the present analysis then is, given a foot template, how the violation of HD(FT) is remedied so that we do not need to postulate an allomorphic syllable template,

hence the elimination of unwarranted heaviness requirements on the syllable template.

Recall that in STS, the headedness requirement on HD(FT) can be respected by an application of INSERT(σ) or COPY(σ). Since the second syllable of disyllabic reduplicants does not contain coda consonants, the headedness requirement should be met by INSERT(σ), rather than by COPY(σ), which, as shown in (9), would lead to total reduplication:



Given that the headedness condition is met by INSERT(σ), (10) compares the unfaithful winning candidate $(\sigma)_{\text{FT}}\text{-pu.kun}$ with another unfaithful candidate $*\text{pu.kun-pu.kun}$ in the first step of the derivation. Since $(\sigma)_{\text{FT}}\text{-pu.kun}$ undergoes INSERT(σ), it violates not only FT-BIN but also HD(σ). The losing candidate $*\text{pu.kun-pu.kun}$ obeys both FT-BIN and HD(σ), but at the expense of filling the foot template by COPY(σ), which violates $*\text{COPY}(\sigma)$. Thus, for $(\sigma)_{\text{FT}}\text{-pu.kun}$ to be more harmonic, $*\text{COPY}(\sigma)$ must outrank both FT-BIN and HD(σ). Note that $(\sigma\sigma)_{\text{FT}}\text{-pu.kun}$ is not a legitimate candidate because it involves two instances of syllable insertion. That is, the derivation from FT-pu.kun to $(\sigma\sigma)_{\text{FT}}\text{-pu.kun}$ violates the *gradualness* requirement of HS.

(10) *1st step of (C)V(C)V reduplication: FT-pu.kun \rightarrow $(\sigma)_{\text{FT}}\text{-pu.kun}$*

$\text{FT} + \begin{matrix} (\sigma & \sigma)_{\text{FT}} \\ \triangle & \triangle \\ \text{pu} & \text{kun} \end{matrix}$	$*\text{COPY}(\sigma)$	FT-BIN	HD(σ)
a. \rightarrow $\begin{matrix} \text{FT} + \begin{matrix} (\sigma & \sigma)_{\text{FT}} \\ \triangle & \triangle \\ \sigma & \text{pu} & \text{kun} \end{matrix} \\ \quad \triangle \quad \triangle \\ \sigma \quad \text{pu} \quad \text{kun} \end{matrix}$		1	1
b. $\begin{matrix} \text{FT} & + & \text{FT} \\ \swarrow & & \searrow \\ \sigma & \sigma & \sigma & \sigma \\ \triangle & \triangle & \triangle & \triangle \\ \text{pu} & \text{kun} & \text{pu} & \text{kun} \end{matrix}$	1W	L	L

Tableau (11), on the other hand, presents a different losing candidate in the first step of the derivation. The form $*\text{FT-pu.kun}$ vacuously obeys HD(σ) because no INSERT(σ) applies. This faithful candidate is ruled out by HD(FT), ranked above HD(σ).

(11) 1st step of (C)V(C)V reduplication: $FT\text{-}pu.kun \rightarrow (\sigma)_{FT}\text{-}pu.kun$

	$FT + (\sigma \quad \sigma)_{FT}$ $\triangle \quad \triangle$ pu kun	HD(FT)	HD(σ)
a. \rightarrow	$FT + (\sigma \quad \sigma)_{FT}$ $\triangle \quad \triangle$ σ pu kun		1
b.	$FT + (\sigma \quad \sigma)_{FT}$ $\triangle \quad \triangle$ pu kun	1W	L

In the second step of the derivation, the winning candidate $(\sigma\sigma)_{FT}\text{-}pu.kun$, as tableau (12) shows, has one more instance of INSERT(σ), hence infringing HD(σ) on more time in favor of FT-BIN. The faithful candidate $*(\sigma)_{FT}\text{-}pu.kun$, in contrast, has only one violation of HD(σ), but fails to meet FT-BIN. Moreover, the unfaithful candidate $*pun\text{-}pu.kun$, though respecting both FT-BIN and HD(σ), does not abide by COPY-LOC(μ), which forbids non-local applications of COPY(μ). Therefore, for $(\sigma\sigma)_{FT}\text{-}pu.kun$ to win out, COPY-LOC(μ) and FT-BIN must dominate HD(σ), which rules out other candidates like $*pu\text{-}pu.kun$ and $*((\mu)_{\sigma})_{FT}\text{-}pu.kun$ that are harmonically bound by $*(\sigma)_{FT}\text{-}pu.kun$ because they additionally violate $*COPY(\mu)$ and HD(μ), respectively.³ Note that $pu_{\mu}k_{\mu}\text{-}pu_{\mu}ku_{\mu}n_{\mu}$ is not a possible candidate. This is so because the input $(\sigma)_{FT}\text{-}pu.kun$ contains an empty syllable whose headedness, under STS, can only be met via COPY(μ) or INSERT(μ). Thus, to derive the candidate $pu_{\mu}k_{\mu}\text{-}pu_{\mu}ku_{\mu}n_{\mu}$ from the input $(\sigma)_{FT}\text{-}pu.kun$ we need two instances of INSERT(μ) and one application of COPY(*seg*). The derivation is illegitimate due to the gradualness requirement of HS.

³ In this stage the candidate $*pu_{\mu}ku_{\mu}\text{-}pu_{\mu}ku_{\mu}n_{\mu}$ is harmonically bound by $*pu_{\mu}\text{-}pu_{\mu}ku_{\mu}n_{\mu}$ because $pu_{\mu}ku_{\mu}$ contains two vowels and is too big a structure for the existing syllable to accommodate. Consequently, the second mora will be left unparsed and FT-BIN is still violated. Moreover, I assume with Hyman (1985), Ito (1986, 1989), and McCarthy et al. (2012) that all segments in a syllable, including onset consonants, are immediate constituents of mora nodes.

(12) 2nd step of (C)V(C)V reduplication: $(\sigma)_{FT}pu.kun \rightarrow (\sigma\sigma)_{FT}pu.kun$

FT + $(\sigma \ \sigma)_{FT}$ $\triangle \ \triangle$ $\sigma \ pu \ kun$	COPY-LOC(μ)	FT-BIN	HD(σ)	*COPY(μ)	HD(μ)
a. \rightarrow $\wedge \ \triangle \ \triangle$ $\sigma \ \sigma \ pu \ kun$			2		
b. $\triangle \ \triangle$ $\sigma \ pu \ kun$		1W	1L		
c. FT + FT \wedge $\sigma \ \sigma \ \sigma$ $\triangle \ \triangle$ $\mu \ pu \ kun$		1W	L		1W
d. FT + FT \wedge $\sigma \ \sigma \ \sigma$ \wedge $\mu \ \underline{\mu} \ \mu \ \mu$ $\triangle \ \underline{\triangle} \ \triangle \ $ $pu \ pu \ ku \ n$		1W	L	1W	
e. FT + FT \wedge $\sigma \ \sigma \ \sigma$ $\wedge \ \ \wedge$ $\mu \ \mu \ \underline{\mu} \ \mu \ \underline{\mu}$ $\triangle \ \ \underline{\triangle} \ \triangle \ $ $pu \ n \ pu \ ku \ n$	1W			1W	

Before moving to the third step of (C)V(C)V reduplication, I need to turn to CVC reduplication for ease of exposition. The situation for the second step of CVC reduplication is a little bit different. Tableau (13) compares the winning candidate $tum.-tum.\beta\partial s$ with the losing one $*(\sigma\sigma)_{FT}tum.\beta\partial s$, both of which abide by FT-BIN. The form $*(\sigma\sigma)_{FT}tum.\beta\partial s$ does not infringe *COPY(μ), of course, but it induces one more violation of HD(σ). The winning candidate, on the other hand, has one more violation of NOCODA (Prince and Smolensky 2004 [1993]). Consequently, HD(σ) must be ranked above both NOCODA and *COPY(μ) for $tum.-tum.\beta\partial s$ to win over $*(\sigma\sigma)_{FT}tum.\beta\partial s$, and the derivation converges in the third step where no more harmonic improvement is possible.

(13) 2nd step of CVC reduplication: $(\sigma)_{FT}\text{-tum.}\beta\text{əs} \rightarrow \text{tum.}\text{-tum.}\beta\text{əs}$

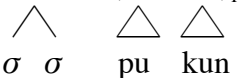
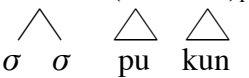
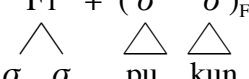
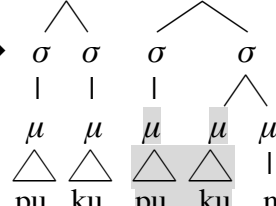

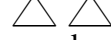

FT + $(\sigma \ \sigma)_{FT}$ $\triangle \ \triangle$ $\sigma \ \text{tum} \ \beta\text{əs}$	FT-BIN	HD(σ)	*COPY(μ)	NOCODA
a. FT + $(\sigma \ \sigma)_{FT}$ $\wedge \ \triangle \ \triangle$ $\sigma \ \sigma \ \text{tum} \ \beta\text{əs}$		2W	L	2L
b. FT + FT \wedge $\rightarrow \ \sigma \ \sigma \ \sigma$ $\wedge \ \wedge \ \wedge$ $\mu \ \mu \ \mu \ \mu \ \mu \ \mu$ $\triangle \ \ \triangle \ \ \triangle \ $ tu m tu m $\beta\text{əs}$ s			1	3

Together, tableaux (10)-(13) illustrate why a foot template can result in heavy syllable reduplication. Given that *COPY(σ) outranks both FT-BIN and HD(σ), the only path for the foot template to eventually appear as disyllabic is via two applications of INSERT(σ). Yet, the constraint *COPY(μ) is ranked below HD(σ). Therefore, in the second step if the first syllable of a stem is CVC (i.e., $tu_\mu m_\mu$ in $tu_\mu m_\mu \beta\text{əs} s_\mu$), COPY(μ) is a more optimal operation to meet FT-BIN because no violation of HD(σ) will be induced. However, if the first syllable of a stem is not heavy (i.e., pu_μ in $pu_\mu ku_\mu n_\mu$), the syllable would contain only one mora. Thus, copying of the only mora will not suffice to satisfy FT-BIN. Note that string-copying the first two moras of the stem $pu_\mu ku_\mu$ would not meet FT-BIN either, because $pu_\mu ku_\mu$ contains two vowels, hence too large to satisfy HD(σ).⁴ Accordingly, HD(σ) will be infringed by one more operation of INSERT(σ) in order to obey FT-BIN, which dominates HD(σ).

Turning back to disyllabic reduplication, Tableau (14) demonstrates that the constraint ranking, HD(σ) » *COPY(μ), also favors $pu.ku\text{-}pu.kun$ in the third step of disyllabic reduplication by ruling out both $*(\sigma\sigma)_{FT}\text{-}pu.kun$ and $*((\mu)_\sigma)_{FT}\text{-}pu.kun$. The derivation of (C)V(C)V reduplication subsequently converges in the fourth step.

⁴ I assume that some undominated constraint penalizes a syllable for accommodating two vowels.

(14) 3rd step of (C)V(C)V reduplication: $(\sigma\sigma)_{FT} \text{-} pu.kun \rightarrow pu.ku \text{-} pu.kun$

	$FT + (\sigma \quad \sigma)_{FT}$  $\sigma \quad \sigma \quad pu \quad kun$	HD(σ)	*COPY(μ)	HD(μ)
a.	$FT + (\sigma \quad \sigma)_{FT}$  $\sigma \quad \sigma \quad pu \quad kun$ μ	1W		1W
b.	$FT + (\sigma \quad \sigma)_{FT}$  $\sigma \quad \sigma \quad pu \quad kun$	2W	L	
c. \rightarrow	$FT + FT$  $\sigma \quad \sigma \quad \sigma \quad \sigma$  $\mu \quad \mu \quad \mu \quad \mu \quad \mu$   $pu \quad ku \quad pu \quad ku \quad n$		1	

In sum, bimoraic reduplication in Kavalan starts as a foot template, whose headedness is met by INSERT(σ) because *COPY(σ) outranks both FT-BIN and HD(σ). The foot template may eventually surface as a CVC syllable if the first syllable of the stem is CVC (i.e., $tu_\mu m_\mu$ in $tu_\mu m_\mu \cdot \beta \partial_\mu s_\mu$). This is so because HD(σ) dominates *COPY(μ) and accordingly COPY(μ) is a more optimal operation to meet FT-BIN. Conversely, if the first syllable of a stem is not heavy (i.e., pu_μ in $pu_\mu \cdot ku_\mu n_\mu$), FT-BIN would not be met by COPY(μ) because the syllable would contain only one mora. As a consequence, in order to obey FT-BIN, HD(σ) will be infringed by one more operation of INSERT(σ), leading to disyllabic reduplication.

4. Heavy Syllable Reduplication: Ilokano Revisited

Since a foot template can eventually be realized as a heavy syllable, it is worth reexamining reduplicative patterns in other languages in which the reduplicant is invariably CVC or CV: and has been analyzed as involving a syllable template. One of such languages is Ilokano, as shown in (15), in which the reduplicant is CVC.

(15) Heavy syllable reduplication in Ilokano (Hayes and Abad 1989: 357-359)

- a. kaldin kal-kaldin ‘goat/pl.’
- b. pu:sa pus-pu:sa ‘cat/pl.’

McCarthy et al. (2012) analyze the pattern as involving a syllable template with an imposed weight condition, which forces the template to ultimately be a heavy

syllable. Yet, how to formulate the weight condition has not been clearly laid out (McCarthy 2012: 197). In fact, no such a requirement should be necessary. The fundamental apparatus of STS alone is capable of deriving the reduplicative pattern in Ilokano. More specifically, considering the fact that coda consonants in Ilokano are moraic (Hayes and Abad 1989: 334), heavy syllable reduplication could be analyzed as involving an application of INSERT(σ) to fill a foot template, followed by two instances of INSERT(μ) in sequence. And the two inserted moras are subsequently filled by an application of COPY(*seg*). In the first step of the derivation the foot template's headedness requirement, because *COPY(σ) outranks HD(σ), is met by INSERT(σ), as shown in (16).

(16) *1st step of heavy syllable reduplication: FT-pu:sa → (σ)_{FT-pu:sa}*

	$\begin{array}{c} \text{FT} + \sigma \quad \sigma \\ \triangle \quad \triangle \\ \text{pu}_{uu} \quad \text{sa} \end{array}$	*COPY(σ)	HD(FT)	HD(σ)	FT-BIN
a.	$\begin{array}{c} \rightarrow \quad \sigma \quad \sigma \\ \triangle \quad \triangle \\ \sigma \quad \text{pu}_{uu} \quad \text{sa} \end{array}$			1	1
b.	$\begin{array}{c} \sigma \quad \sigma \\ \triangle \quad \triangle \\ \text{pu}_{uu} \quad \text{sa} \end{array}$		1W	L	1
c.	$\begin{array}{c} \sigma \quad \sigma \\ \triangle \quad \triangle \quad \triangle \\ \text{pu}_{uu}\text{sa} \quad \text{pu}_{uu} \quad \text{sa} \end{array}$	1W		L	L

In the second step, illustrated in (17), fulfillment of the inserted syllable takes priority than that of foot binarity because HD(σ) dominates FT-BIN. The headedness of the inserted syllable is met by INSERT(μ) for COPY(μ) outranks HD(μ).

(17) 2nd step of heavy syllable reduplication: $(\sigma)_{FT}pu:sa \rightarrow ((\mu)_{\sigma})_{FT}pu:sa$

	FT + σ σ \triangle \triangle σ pu _{uu} sa	HD(σ)	FT-BIN	*COPY(μ)	HD(μ)
a. \rightarrow	FT + σ σ \triangle \triangle σ pu _{μ} sa μ		1		1
b.	FT + σ σ \triangle \triangle σ pu _{uu} sa	1W	1		L
c.	FT + σ σ \wedge \triangle \triangle σ σ pu _{uu} sa	2W	L		L
d.	FT + σ σ \wedge \triangle σ μ μ sa \wedge / μ p u \triangle pu		1	1W	L

In the 3rd step harmonic improvement is achieved via one more application of INSERT(μ), violating of HD(μ) in favor of FT-BIN and HD(σ). This is shown in tableau (18).

(18) 3rd step of heavy syllable reduplication: $((\mu)_\sigma)_{FT}pu:sa \rightarrow ((\mu\mu)_\sigma)_{FT}pu:sa$

	FT + σ σ	HD(σ)	FT-BIN	*COPY(μ)	HD(μ)
	$\begin{array}{c} \quad \triangle \quad \triangle \\ \sigma \quad pu_{\mu\mu} \quad sa \\ \\ \mu \end{array}$				
a. \rightarrow	$\begin{array}{c} FT + \sigma \quad \sigma \\ \quad \triangle \quad \triangle \\ \sigma \quad pu_{\mu\mu} \quad sa \\ \wedge \\ \mu \quad \mu \end{array}$				2
b.	$\begin{array}{c} FT + \sigma \quad \sigma \\ \quad \triangle \quad \triangle \\ \sigma \quad pu_{\mu\mu} \quad sa \\ \\ \mu \end{array}$		1W		1L
c.	$\begin{array}{c} FT + \sigma \quad \sigma \\ \wedge \quad \triangle \quad \triangle \\ \sigma \quad \sigma \quad pu_{\mu\mu} \quad sa \\ \\ \mu \end{array}$	1W			1L

Since both *COPY(seg) and NOCODA are ranked the lowest, the fourth step proceeds with an application of COPY(seg), which copies a CVC string to satisfy HD(μ) without additional infringement of NO-LONG-V (Rosenthal 1994),⁵ as shown in (19)-(20). The derivation then converges in the fifth step.

⁵ I assume that long vowels in the base are protected by higher ranked faithfulness constraints.

(19) 4th step of heavy syllable reduplication: $((\mu\mu)_\sigma)_{FT}pu:sa \rightarrow pus.-pu:sa$

$FT + \sigma \quad \sigma$ $\mu \quad \mu$ $pu_{\mu\mu} \quad sa$	HD(μ)	NO-LONG-V	NOCODA	*COPY(<i>seg</i>)
<p>a. \rightarrow</p> $FT +$ $\mu \quad \mu \quad \mu \quad \mu \quad \mu$ $\triangle \quad \quad \triangle \quad / \quad \triangle$ $pu \quad s \quad p \quad u \quad sa$		1	1	1
<p>b.</p> $FT + \sigma \quad \sigma$ $\mu \quad \mu$ $pu_{\mu\mu} \quad sa$	2W	1	L	L
<p>c.</p> $FT +$ $\mu \quad \mu \quad \mu \quad \mu \quad \mu$ $\triangle \quad / \quad \triangle \quad / \quad \triangle$ $p \quad u \quad p \quad u \quad sa$		2W	L	1
<p>d.</p> $FT +$ $\mu \quad \mu \quad \mu \quad \mu \quad \mu$ $\triangle \quad \quad \triangle \quad / \quad \triangle$ $pu \quad p \quad u \quad sa$	1W	1	L	1

(20) 4th step of heavy syllable reduplication: $((\mu\mu)_\sigma)_{FT}\text{-kal.dij} \rightarrow \text{kal.}\text{-kal.dij}$

$\begin{array}{c} \text{FT} + \sigma \quad \sigma \\ \quad \triangle \quad \triangle \\ \sigma \quad \text{kal} \quad \text{dij} \\ \wedge \\ \mu \quad \mu \end{array}$	HD(μ)	NO-LONG-V	NOCODA	*COPY(<i>seg</i>)
<p>a. \rightarrow</p> $\begin{array}{c} \text{FT} + \\ \quad \sigma \quad \sigma \quad \sigma \\ \wedge \quad \wedge \quad \triangle \\ \mu \quad \mu \quad \mu \quad \mu \quad \text{dij} \\ \triangle \quad \quad \triangle \quad \\ \text{ka} \quad \text{l} \quad \text{ka} \quad \text{l} \end{array}$			3	1
<p>b.</p> $\begin{array}{c} \text{FT} + \sigma \quad \sigma \\ \quad \triangle \quad \triangle \\ \sigma \quad \text{kal} \quad \text{dij} \\ \wedge \\ \mu \quad \mu \end{array}$	2W		2L	L
<p>c.</p> $\begin{array}{c} \text{FT} + \\ \quad \sigma \quad \sigma \quad \sigma \\ \wedge \quad \wedge \quad \triangle \\ \mu \quad \mu \quad \mu \quad \mu \quad \text{dij} \\ \wedge \quad \wedge \quad \triangle \quad \\ \text{k} \quad \text{a} \quad \text{ka} \quad \text{l} \end{array}$		1W	2L	1
<p>d.</p> $\begin{array}{c} \text{FT} + \\ \quad \sigma \quad \sigma \quad \sigma \\ \wedge \quad \wedge \quad \triangle \\ \mu \quad \mu \quad \mu \quad \mu \quad \text{dij} \\ \triangle \quad \triangle \quad \\ \text{ka} \quad \text{ka} \quad \text{l} \end{array}$	1W		2L	1

In short, contra McCarthy et al. (2012: 197), the present analysis shows that heavy syllable reduplication in Ilokano is derivable in STS if we assume that the template involved is a foot. It follows that no heaviness requirement on a syllable template is necessary.

5. Conclusion

This article has shown that variable reduplication in Kavalan, as well as heavy syllable reduplication in Ilokano, can be tackled in Serial Template Satisfaction without stipulations of any weight conditions or multiple allomorphic templates. The present analysis strongly suggests that whenever a monosyllabic reduplicant is required to be heavy, the heaviness requirement should be reduced to the basic machinery of STS and the language's phonology. The abandonment of weight requirements naturally leads to the conclusion that templates below the prosodic

word level are just FT and σ , but never σ_μ or $\sigma_{\mu\mu}$.⁶ Moreover, by eliminating weight requirements, we also dispense with $RED=\mu\mu$ as a constraint on the reduplicant. Crucially, since weight conditions should not exist, it would be no longer necessary to assume allomorphic templates to account for variable reduplication. If the reduplicant in a language is invariably bimoraic, as in Ilokano, or is bimoraic in general, wider distribution, as in Kavalan, then the template should be a foot, regardless of whether the reduplicant is disyllabic or a heavy syllable. Ideally, every variable pattern should emerge as a consequence of constraint interaction in STS.

References

- Blevins, Juliette. 1996. Mokilese reduplication. *Linguistic Inquiry* 27: 523-530.
- Chang, Yung-li. 2000. *A reference grammar of Kavalan*. Taipei: Yuan-liu.
- Crowhurst, Megan. 2004. Mora alignment. *Natural Language and Linguistic Theory* 22: 127-177.
- Elfner, Emily. 2009. Syllabification and stress-epenthesis interactions in Harmonic Serialism. Ms., University of Massachusetts, Amherst, Ma.
- Hayes, Bruce and May Abad. 1989. Reduplication and syllabification in Ilokano. *Lingua* 77: 331-374.
- Hyman, Larry. 1985. *A theory of phonological weight*. Dordrecht: Foris.
- Ito, Junko. 1986. Syllable theory in prosodic phonology. Doctoral dissertation, University of Massachusetts, Amherst.
- Ito, Junko. 1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7: 217-259.
- Kimper, Wendell. 2011. Locality and globality in phonological variation. *Natural Language and Linguistic Theory* 29.2: 423-465.
- Lin, Hui-shan. 2012. Variations in Kavalan reduplication. *Language and Linguistics* 13.6: 1051-1093.
- Lin, Kuo-Chiao. 2014. Gradual epenthesis: Echo vowels in Austronesian languages. *Oceanic Linguistics* 53.2: 443-466.
- Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13: 483-545.
- McCarthy, John. 2000. Harmonic Serialism and parallelism. *Proceedings of NELS 30*, 501-524. *Proceedings of NELS 19*, 191-206. GLSA, University of Massachusetts, Amherst.
- McCarthy, John. 2008a. The gradual path to cluster simplification. *Phonology* 25: 271-319.
- McCarthy, John. 2008b. The serial interaction of stress and syncope. *Natural Language and Linguistic Theory* 26: 499-546.
- McCarthy, John. 2010. An introduction to Harmonic Serialism. *Language and*

⁶ I set aside the question as to whether there should be mora templates. Note, however, that since templates are defined with reference to prosodic tiers, weight requirements should be interpreted as involving two concatenative mora templates.

- Linguistics Compass* 4.10: 1001-1018.
- McCarthy, John, Wendell Kimper and Kevin Mullin. 2012. Reduplication in Harmonic Serialism. *Morphology* 22: 173-232.
- McCarthy, John and Alan Prince. 1993b. Prosodic morphology I: Constraint interaction and satisfaction. Ms., University of Massachusetts, Amherst, Ma., and Rutgers University, New Brunswick, Nj.
- McCarthy, John and Alan Prince. 1995. *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*, 249-384. University of Massachusetts, Amherst, Ma.
- McCarthy, John and Alan Prince. 1996 [1986]. *Prosodic morphology 1986*. Ms., Rutgers University Center for Cognitive Science, New Brunswick, Nj.
- McCarthy, John and Alan Prince. 1999. Faithfulness and identity in prosodic morphology. *The prosody-morphology interface*, eds. Rene Kager, Harry van der Hulst and Wim Zonneveld, 218-309. Cambridge: Cambridge University Press.
- Prince, Alan. 1980. A metrical theory for Estonian quantity. *Linguistic Inquiry* 11: 511-562.
- Prince, Alan and Paul Smolensky. 2004 [1993]. *Optimality Theory: Constraint interaction in Generative Grammar*. Ms., Rutgers University, New Brunswick, Nj.
- Pruitt, Kathryn. 2010. Serialism and locality in constraint-based metrical parsing. *Phonology* 27.3: 481-526.
- Rosenthal, Sam. 1994. Vowel glide alternation in a theory of constraint interaction. Doctoral dissertation, University of Massachusetts, Amherst.
- Selkirk, Elisabeth. 1995. The prosodic structure of function words. *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*, 439-470. University of Massachusetts, Amherst, Ma.
- Thurgood, Elzbieta. 1997. Bontok reduplication and prosodic templates. *Oceanic Linguistics* 36: 135-148.
- Wolf, Matthew. 2008. Optimal interleaving: Serial phonology-morphology interaction in a constraint-based model. Doctoral dissertation, University of Massachusetts, Amherst.

THE PROCESSING OF LONG-DISTANCE DEPENDENCIES IN
NIUEAN*

Nicholas Longenbaugh Maria Polinsky
MIT University of Maryland
nslonge@mit.edu polinsky@umd.edu

It is well documented that nominative-accusative alignment coincides with a strong *subject-preference* in long-distance dependency formation, both in terms of processing (in particular, subject gaps in relative clauses are processed more easily than other types of gaps) and accessibility for extraction (there appears to be an implicational universal that the availability of relativization with a gap entails the availability of subject relativization with a gap (Keenan and Comrie 1977)). This subject preference does not, however, carry over uniformly to languages with ergative-absolutive alignment; in some morphologically ergative languages, the ergative subject cannot extract with a gap at the extraction site, a phenomenon termed *syntactic ergativity*. In this paper, we explore the viability of a processing-based explanation of syntactic ergativity. Much as has been proposed for various island phenomena (Kluender 1998, Kluender 2004), the extraction of ergative arguments may simply be more taxing on the parser than corresponding absolutive extraction. If this is true, following Hawkins (2004, 2014), syntactically ergative languages could then be taken to differ from their morphologically ergative counterparts in their tolerance for difficult structure, eliminating the less efficient, more difficult ergative extraction. To test this account, we explore the processing of relative clauses in Niuean, a morphologically, but not syntactically, ergative language. Niuean is an ideal test case, given that it is closely related to the syntactically ergative Tongan, and thus might be expected to show an obvious bias against ergative extraction (a bias that Tongan turns into a categorical restriction). We present novel experimental data showing that ergative subject gaps in Niuean RCs do not impose any additional processing difficulty as compared to the processing of absolutive object gaps, thus calling into question the viability of the processing account in this domain.

*This work was funded by NSF grants BCS-1144223, BCS-1342688, BCS-1414318, and SMA-1429961 to Maria Polinsky. We would like to thank Kara Tukuitoga, Granby and Pat Siakimoto, Tom Etuata, Mele Nemaia, and Pefi King for sharing their language with us. We are grateful to Jason Brown and Miriam Meyerhoff for their support and helpful comments, to Annie Gagliardi and Zuzanna Fuchs for technical assistance, and to the audiences at the University of Auckland and at AFLA-22 in Montreal for their input. All errors are our responsibility.

1. Introduction

Subjects are special. Since at least the work of Keenan and Comrie (1977, 1979), it has been known that there is a pervasive cross-linguistic tendency to prefer them over other arguments and adjuncts from the perspective of a number of syntactic operations. In this paper, we take up one facet of this phenomenon: the fact that subject gaps in long-distance dependencies are interpreted more accurately and faster than other types of gaps (Kwon et al. 2013). This subject preference is captured in Keenan and Comrie's (1977) *Accessibility Hierarchy*, which rates grammatical function based on the relative ease with which they may form relative clauses and possibly, other A'-dependencies.

- (1) Accessibility hierarchy:
Subject >> object >> indirect object >> oblique object >> possessor >>
object of comparison
(Keenan and Comrie 1977)

Of course, *subject* is itself a notion that encapsulates a number of variable and often unrelated properties, so that we might question exactly what is about *subjecthood* that leads to this effect.

In general terms, there are any number of factors that might be implicated in the subject preference, including: the thematic status of the subject (subjects tend to be agentive); the status of the gap associated with subject dependencies (subject and non-subject gaps might differ with respect to whether they are traces or bound pronouns); the morphological or abstract case associated with the gap (nominative vs. non-nominative, ergative vs. non-ergative); the general structural properties associated with either the subject-gap itself or the dependency chain more generally, and other factors. Even assuming all such properties play a role, there is reason to suspect that morphological case has a special status in the determination of the subject preference. Indeed, most of the languages included in Keenan and Comrie's (1977) initial survey, and that have been studied in this context since, have been *nominative-accusative* languages, where the grammatical function and (nominative) morphological case align in most clauses. If the basis of consideration is widened to include languages exhibiting an *ergative-absolutive* alignment, the neat pattern encapsulated in the accessibility hierarchy appears to break down. So-called *ergative* languages morphologically distinguish subjects of transitive verb, which surface with *ergative* case, to subjects of intransitive verbs and objects of transitive verbs, which surface with *absolutive* case.

(2) Accusative alignment:

Ergative Alignment

Nominative	Accusative	Ergative	Absolutive
-transitive subject	-transitive object	-transitive subject	-transitive object
-intransitive subject			-intransitive subject

Ergative subjects seem to have the typical properties we associate with subject (Anderson 1976, and much subsequent literature) and thus behave identically to nominative subjects with respect to binding (they bind structurally lower arguments but cannot be bound by them), imperative formation (they appear as the addressee of an imperative), accessibility to control and raising, and preferential status in maintaining coreference across clauses. However, despite their subject-like behavior, in a large number of languages with ergative-absolutive alignment, the ergative subject cannot be part of a long-distance dependency leaving a gap at the base position. Absolutive arguments, subject or object, display no such restrictions, freely participating in A'-dependency formation with a gap. This restriction on the extraction of the ergative is known as *syntactic ergativity*.

One can imagine approaching the problem of syntactic ergativity from at least two perspectives, roughly mirroring extant approaches to other restrictions on long-distance dependencies. First, we could seek to attribute the unavailability of ergative extraction to independently motivated syntactic principles. Under this style of approach, the syntax of languages with syntactic ergativity differs in some meaningful way from accusative languages and from morphologically ergative languages without syntactic ergativity (for varieties of this idea, see Aldridge 2008, Legate 2008, Polinsky et al. 2012, Coon et al. 2015). We'll refer to these analyses broadly as instantiating the *syntactic account*. An alternative approach, which we will refer to as the *processing account*, is to explain syntactic ergativity through processing constraints, much as has been proposed for various island phenomena (Kluender 1998, Kluender 2004) and accessibility hierarchies. With respect to the latter, Hawkins (2014) writes:

“that there are several hierarchies of *wh*-movement and relativization, with their systematic cut-off points in different languages and alternations between gaps and resumptive pronouns in easier and more difficult environments respectively, [can] be readily explained in terms of increasing processing complexity” (Hawkins 2014: 81)

Under such a view, extraction of ergative arguments would be held to be more taxing on the parser than corresponding absolutive extraction. Following the general conception developed by Hawkins (2004, 2014), syntactically ergative languages would then differ from their merely morphologically ergative counterparts in having a different “systematic cut off point,” with the syntactically ergative languages requiring a resumptive pronoun in the ergative gap site.

Given that processing explanations have been successful in accounting for preferences in categorization or word order principles (Hawkins 2014), a processing approach to syntactic ergativity warrants serious consideration. In order to evaluate the processing account, it would be desirable to explore a pair of similar languages, where one has syntactic ergativity and the other does not. If the language without syntactic ergativity in the comparison set can be shown to impose a heavier processing load on long-distance dependencies with an ergative gap as compared to those with absolutive gaps, such a difference would lend credence to the notion that the related language – one with syntactic ergativity – simply eliminated the less efficient, more difficult structure. Polynesian languages offer an ideal test case in this regard. The closely related Tongan and Niuean are very similar in many ways, although they differ with respect to syntactic ergativity: Tongan is syntactically ergative and Niuean is not.

In what follows, we present novel experimental data on the processing of gaps in Niuean relative clauses. To anticipate our conclusions, we show that the processing of ergative gaps in Niuean does not impose any additional processing difficulty as compared to the processing of absolutive object gaps, thus calling into question the applicability of the processing account to syntactic ergativity.

The remainder of this paper is structured as follows. In Section 2 we sketch the basics of syntactic ergativity, as well as our motivation for comparing Tongan and Niuean. Section 3 comprises a presentation of our experimental methods, materials, and results. We discuss these results in Section 4, both for accounts of syntactic ergativity and for the related notion of the subject preference more generally. We offer a conclusion in Section 5.

2. Syntactic ergativity

2.1. The phenomenon

For the purposes of this paper, we will adopt the relatively restricted definition of syntactic ergativity provided in 3, setting aside the variety of related phenomena that have variously been grouped under the same header.

(3) *Syntactic Ergativity:*

The grouping together of intransitive subject and transitive object (absolutive arguments), to the exclusion of transitive subjects (ergatives), with respect to accessibility to A'-dependency formation.

(Polinsky to appear a)

For the most part our investigation will be further limited to discussion of relative clause formation. There are several reasons for that. As we have already mentioned, the accessibility hierarchy (see 1) was explicitly formulated to reflect the ease of relative-clause formation, and syntactic ergativity emerges as a challenge to this hierarchy. Next, A'-dependencies in the two languages under discussion here, Ton-

cessibility hierarchy thus appears to accurately describe the behavior of nominative-accusative languages, where we never find a broad-based restriction on the long-distance extraction of subjects if such a restriction does not also obtain for objects and other DPs.

The constellation of facts surrounding syntactic ergativity lends at least initial credibility to processing-based accounts. Because a large number morphologically ergative languages exhibit syntactic ergativity, morphologically but not syntactically ergative languages take on an immediate *outlier* status. If ergative gaps in these outlier languages can be shown to be more difficult to process than absolutive gaps, then variability in terms of the presence of syntactic ergativity, and the behavior of the outlier languages could be reduced to differential threshold for complex data sets. That languages differ in their tolerance for complex structure is well known, with one illustrative example being the acceptability of certain crossed-dependencies in the Scandinavian languages. Norwegian tolerates crossed-dependencies – which are notoriously complex, both formally and from a processing perspective (see Fodor 1978; Engdahl 1979; Pesetsky 1982; Creider 1986; Miller 1991) – formed via extraction from embedded questions without recourse to resumptive pronouns. In Swedish, on the other hand, a resumptive pronoun must appear in all gap sites except the hierarchically lowest in all instances of extraction leading to crossed-dependencies, but not if the corresponding dependencies are nested. Engdahl (1982) and Asudeh (2004) argue convincingly that the resumptive pronouns that appear in these cases serve as processing aides in Swedish, so that Norwegian, in not requiring such pronouns, can be seen to be more tolerant of complex structure parsing.

- (6) a. [den här presenten]₁ kan du säkert aldrig komma på vem₂ jag
 this here gift can 2.SG surely never arrive at who 1.SG
 fick *(den₁) av ___₂
 got 3.SG.RP from
 ‘this gift that you’ll never guess who I got (it) from’ *Swedish*
- b. [denne gaven her]₁ vil du ikke gjette hvem₂ jeg fikk (den₁)
 this gift here will 2.SG not guess who 1.SG got 3.SG.RP
 fra ___₂
 from
 ‘this gift that you’ll never guess who I got (it) from’ *Norwegian*
 (Maling and Zaenen 1982: 236)

Before examining a processing account, we will briefly survey the alternative accounts, which as we mentioned above, appeal to syntactic principles rather than processing.

2.2. Existing accounts

In spite of the intuitive appeal of a processing-based account, until recently, most work on syntactic ergativity has taken the perspective that the phenomenon is syntactic in nature. Bittner (1994) and Bittner and Hale (1996) develop an account that hinges on the assignment of absolutive case, which they assume takes place in a vP-external A'-position. Under the further assumption that A'-movement for case purposes takes precedent over all other types of A'-movement, the derivation will always prefer movement of absolutive to its vP-external case position over other A'-dependency formation. Consequently, A'-extraction of the vP-internal ergative will always be blocked if A'-movement is successive-cyclic, given that the absolutive occupies the relevant left-peripheral A'-landing site (see Polinsky to appear for more discussion). Languages without syntactic ergativity might then differ in permitting multiple A'-specifiers or in otherwise permitting A'-extraction past an occupied A'-position.

A related family of approaches also relies on movement of the absolutive to a position c-commanding the ergative as a means of deriving syntactic ergativity. The work of Legate (2006, 2008) and Aldridge (2004, 2008) best represents this take on the phenomenon. These authors assume that, in at least some ergative languages, absolutive case is assigned differently in transitive and intransitive clauses. They encode this differential behavior in the v head, so that transitive v assigns absolutive case, while intransitive v does not. Intransitive clauses thus involve movement of the sole argument to spec,TP for case purposes. In transitive examples, absolutive case is predicated on EPP driven movement of the object to an outer specifier of vP. This movement places the absolutive in a hierarchically superior position to the ergative, which presumably occupies either an inner specifier of vP or some other vP-internal position. With this structural configuration, A'-movement of the ergative can be blocked by the *attract closest principle* (see 7) under the assumption that the absolutive is also a suitable target for movement. These authors accomplish this by positing that A'-movement in syntactically ergative languages is driven by a probing EPP feature, for which presumably any DP is a suitable goal. A'-movement in non-syntactically ergative languages might then be said to involve, for example, [wh]-features, so that the ergative and absolutive need not both be potential candidates for a given movement operation.

(7) *Attract closest principle:*

A head which attracts a given kind of constituent attracts the closest constituent of the relevant kind

(Richards 1997)

Coon and Preminger (2012) have proposed a related approach grounded in phase-theory (Chomsky 2001, 2008). Drawing on the behavior of a number of Mayan languages, they contend that at least in this family, ergative languages can be classified into two groups based on whether they license absolutive case vP-externally, the

so-called *high*-absolutive languages, or vP internally, the so-called *low*-absolutive languages. It follows that the absolutive argument must move out of the vP for case purposes in high-absolutive languages, again leading to a configuration where the absolutive c-commands the ergative. Syntactic ergativity follows under this account from the proposal that the transitive v is a phase head: by assumption, phases have exactly one *escape hatch* through which arguments may move, so that movement of the absolutive to its vP-external case position effectively blocks any other extraction, be it A or A', from the vP.

All of the above analyses share a reliance on the properties of the absolutive argument in deriving syntactic ergativity. An alternative, taken up by at least Stepanov (2004), Markman and Grashchenkov (2012), and Polinsky (to appear b), is to treat the ergative argument itself as the locus of syntactic ergativity. These authors all propose to derive this behavior by treating the ergative argument in syntactically ergative languages as a PP. For Stepanov (2004), ergative subjects, which are PPs, are adjoined at a late stage in the derivation, crucially after the action of cyclic syntactic rules like A'-movement. This account predicts that syntactically ergative languages should also lack agreement with the ergative, which appears in the derivation after all agreement operations have taken place. There do, however, appear to be syntactic ergative languages which show genuine agreement with the ergative: Halkomelem (Gerds 1988; Wiltschko 2006), Abkhaz (Hewitt 1979), Adyghe (Testelefs et al. 2009), and Chukchi (Bobaljik and Branigan 2006; Polinsky to appear b). To avoid this pitfall, Markman and Grashchenkov (2012) and Polinsky (to appear b) reject the late-merge hypothesis and allow PPs to enter into agreement relationships with verbal heads. For Polinsky (to appear b), syntactic ergativity can then be seen to follow from the PP status of the ergative under the assumption that syntactically ergative languages are exactly those that do not have P-stranding or pied-piping.

Given the somewhat fragmented state that characterizes the space of syntactic accounts, it is appealing to think that a single, uniform processing explanation may supplant these accounts. To this end, as mentioned in the introduction, it would be ideal to find a pair of similar, closely related languages, where one but not the other is syntactically ergative. If the main idea underlying the processing account, that A'-dependencies involving ergative arguments incur a higher processing cost than equivalent dependencies involving absolutive arguments, is correct, we should expect to be able to observe it in the language without syntactic ergativity. Tongan and Niuean, members of the Tongic branch of Western Polynesian group, present an excellent pair for exactly this kind of investigation.

2.3. Setting the stage: Tongan and Niuean

As example 4 demonstrates, Tongan exhibits syntactic ergativity. Niuean, on the other hand, *requires* a gap at the extraction site for relative clauses formed of both ergative and absolutive arguments (Seiter 1980).²

²Note the near synonymy of the examples from Tongan and Niuean.

- (8) a. e *fifine*₁ [ne fano (*a ia₁) ki Toga] *Absolutive subject gap*
 ABS woman NFT go ABS 3SG OBL to Tonga
 ‘the woman who went to Tonga’
- b. e *fifine*₁ [ne ofaofa e Sione (*a ia₁)] *Absolutive object gap*
 ABS woman NFT love ERG Sione ABS 3.SG
 ‘the woman who Sione loves’
- c. e *fifine*₁ [ne ofaofa (*e ia₁) a Sione] *Ergative gap*
 ABS woman NFT love ERG 3.SG ABS Sione
 ‘the woman who loves Sione’

Niuean is therefore an ideal test case for evaluating the processing account: if it can be shown that its ergative gaps are significantly harder to process than absolutive gaps, it would lend credence to the notion that the closely related Tongan has removed the difficult structure altogether. In addition to offering us a valuable opportunity to test the processing explanation of syntactic ergativity, Niuean warrants a processing examination in its own right; our knowledge of the processing of long-distance dependencies is generally limited to large, thriving languages (English, German, Japanese, etc.), and understanding the processing of such dependencies in a small, endangered language is valuable in and of itself (see Sperlich 1995; Sperlich 2005 for discussion of the endangered status of Niuean).

Let us review some of the key properties of Niuean.³ Reports vary, but there are an estimated 7,000 speakers of Niuean, of whom perhaps 2,000 or 3,000 know the language fluently and use it on a daily basis. Most fluent speakers are over 40 years old, so the language is critically endangered (see also Sperlich 1995). The majority of Niuean speakers reside in Auckland, New Zealand; the number of speakers on the island of Niue has been declining precipitously, and currently there may be a thousand speakers left there. Niuean has a modest written tradition, with major literary sources confined to the Bible and various children’s and instructional books created by the New Zealand Ministry of Education.

Niuean is universally verb- and predicate-initial, with alternate VSOX and VOSX word order (Massam 2000, 2001). The language exhibits ergative alignment in all clauses, marking this alignment with overt case morphology; there is no agreement morphology⁴. Case is marked by particles that precede the relevant noun phrase, and these particles are divided into two classes: pronominal (used with pronouns and proper nouns) and common.

³Because our study involves the processing of A’-dependencies in Niuean, not Tongan, we focus here on the essential properties of Niuean only, although when relevant we point out similarities between the languages.

⁴The use of reduplicated or suppletive forms with some verbs that denote an event with plural subject is not agreement (pace Seiter 1980) but the marking of pluractionality Haji-Abdolhosseini et al. (2002).

- (12) Relativization with a resumptive pronoun
- a. e tau tagata ne age e ia *(ki ai) e motoka
 ABS PL person NFT give ERG 3.SG to 3.PL.RP ABS car
 ‘the people he gave the car to’
- b. e taga ne tuku *(ai) e ia e uga
 ABS bag NFT put 3.SG.RP ERG 3.SG ABS crab
 ‘the bag in which he put the coconut crab’
 (Seiter 1980: 94)

Niuean RCs do not appear to involve head-raising (Bhatt 2002, Hulsey and Sauerland 2006), as they fail to permit idiom chunk reconstruction (see 13) or relative-clause-internal scope for quantified head-nouns (see 14).

- (13)a. Kua kai e ia haana a kokō.
 PERF eat ERG 3.SG 3.SG.POSS DET vomit
 ‘(S)he is a hypocrite’ (lit. (s)he ate her/his vomit)
- b. e kokō₁ [ne kai e ia ___₁]
 DET vomit COMP eat ERG 3.SG
 ‘the vomit that (s)he ate’, not ‘her/his hypocrisy’
- (14)a. Kua lagomatai he tolu ekekafo e tokoua faiaoga.
 PERF help ERG three doctors ABS two teachers
 ‘Three doctors helped two teachers.’
 3>>2, 2>>3
- b. e tolu ekekafo₁ [ne lagomatai ___₁ e tokoua faiaoga]
 DET three doctor COMP help ABS two teachers
 ‘the three doctors that helped two teachers...’
 3>>2, ?*2>>3
- c. e tokoua faiaoga₁ [ne lagomatai he tolu ekekafo ___₁]
 DET two teachers COMP help ERG three doctors ___₁
 ‘the three teachers that two doctors helped...’
 2>>3, *3>>2

These facts suggest an operator-based analysis of Niuean relative clauses is most appropriate. Whether the relevant operator moves to the left periphery, or is base-generated there and unselectively binds a null *pro*, as has been recently suggested for a number of related A'-dependency chains in Tongan (Polinsky to appear b) and various Formosan languages (Chen & Fukuda, this volume), is difficult to determine. For present purposes, what matters is that relativization of ergative and absolutive arguments obtains equally with a gap, and we leave the syntactic category of that gap (trace or *pro*) for further investigation (see 15).

- (15) NP [_{CP} Op₁ ... t₁/pro₁]

To determine the frequency of absolutive and ergative gaps in relative clauses, we completed a manual corpus survey conducted on 2,300 clauses from an interview text (courtesy Donna Sparks) and children’s books from which we extracted relative clauses. The distribution of ergative, absolutive subject, and absolutive object gaps in the relative clauses in this mini-corpus was roughly as in 16. We provide comparable statistics from English for reference, based on the data in (Gordon and Hendrick 2005), and from Avar (Polinsky et al. 2012).

- (16) Niuean, English, and Avar relative clause distribution (percentages rounded off to integers)

	Intransitive subject gap	Transitive subject gap	Object gap
<i>Niuean</i>	38%	31%	31%
<i>English</i>	34%	32%	34%
<i>Avar</i>	43%	26%	31%

The distribution of the three gap types that are of interest to us is comparable within Niuean and also comparable to other languages. This is a pleasing fact; if the number of a particular gap type were found to be significantly greater than the number of other gap types, our results could be explained simply in terms of frequency or familiarity of the structures. The distribution established here indicates that frequency or familiarity should not be cause for concern. With this information in place, we now turn to outlining our experimental methods and procedures.

3. Processing study: Niuean

To test the predictions of the processing account, we examined the relative ease of processing of absolutive subject gaps, absolutive object gaps, and ergative gaps in Niuean.

3.1. Participants, materials, procedure

Forty seven speakers of Niuean, recruited in and around Auckland, New Zealand, participated in this study. Data from 6 subjects were excluded due to extremely long response times (>10s per question) and high error rates. The remaining forty one subjects, 27 female, had a mean age of 42;8. All the subjects were Niuean native speakers, the majority of them born in Niue, who continue to use Niuean on a daily basis. Subjects were screened for fluency with the aid of two native speaker consultants who travelled with the authors to identify subjects and who assisted in the experiment. All subjects also spoke English, with a minority bilingual from birth.⁵

⁵At least two studies have shown that bilingual speakers respond differently than their monolingual counterparts when exposed to identical stimuli (Zawiszewski et al. 2011, Clemens

Because of the extremely modest reading tradition in Niuean (see our remarks above), we opted to conduct an auditory picture-matching study that required no familiarity with written Niuean. The utility of picture-matching studies has been demonstrated in a similar experiment conducted with Mayan speakers in Guatemala and Mexico (Clemens et al. 2015).⁶ In a within-subjects design, subjects saw pictures on a computer screen depicting three participants, then heard a question related to the picture that asked them to identify one of the participants.

The images were constructed according to the following general paradigm. All images depicted three participants X_1 , Y , X_2 , with X_1 and X_2 from the same general category, disparate from the category of Y . Thus the X participants were, variously, boys, ships, horses, dogs, etc., and the corresponding Y participants were girls, sharks, cows, cats, etc., respectively. Participants were selected so that each image contained either all human participants, all animal participants, or all inanimate participants. The participants depicted in the images were normed with three native speakers to ensure that they were culturally salient and easily recognizable. Participants were arranged sequentially along the horizontal axis, with X_1 to the left of Y to the left of X_2 . In Class A (intransitive) images, X_1 and Y both carry out the same intransitive action (sleeping, laughing, crying, etc.), while X_2 carries out a distinct intransitive action. In the Class B (transitive) images, X_1 acts on Y , which acts on X_2 , where the action carried out is the same for both X_1 and Y . Sample images appear below. Each of the twenty five images was computationally inverted about the vertical axis as well, producing a total of fifty images.



(a) Class A (X_1 , Y are sweeping, X_2 is smiling) (b) Class B (X_1 is licking Y is licking X_2)

Figure 1: Experimental images

et al. 2015). Given that all of our subjects were bilingual and there are virtually no monolingual speakers of Niuean left, we have no way of confirming or controlling for aberrations related to this.

⁶Clemens et al. (2015) present a direct comparison between picture-matching and self-paced reading which presents the proof of concept showing that these two methodologies yield comparable results. See also our discussion below, of the two methodologies used in Avar.

Auditory stimuli consisted of questions asking subjects in the experiment to identify one of the participants in the corresponding image. All questions made use of a relative clause whose head noun referred to participant in the experimental images. The relative clauses varied according to transitivity and gap site, with each of transitive subject gap, intransitive subject gap, and transitive object gap attested. The duration of the experimental question was made comparable for transitive and intransitive stimuli; in order to make intransitive stimuli comparable to the transitive ones in duration, the relative clauses inside those included a PP or an adverb, as in example (19a).

Each Class A (intransitive) image was associated with questions of the following form, where the categories of participants are X, Y, and the relevant intransitive action is IV.

- (17)a. Where is the X that is IVing Adv/PP?
- b. Where is the X that is not IVing Adv/PP?

Likewise, each Class B (transitive) image was associated with questions of the following form, where the categories of participants are X, Y, and the relevant transitive action is TV.

- (18)a. Where is the X that is TVing the Y?
- b. Where is the X that the Y is TVing?
- c. Where is the Y that is TVing the X?
- d. Where is the Y that the X is TVing?

Twenty four stimuli were created along these lines, with eight questions per gap type (intransitive subject, transitive object, transitive subject). Of these eight questions per gap type, three involved pictures with human participants, three with animal participants, and two with inanimate participants. In addition to these stimuli, which always question one of the X participants, we also constructed twenty four stimuli questioning entity Y. Because such questions do not require resolution of the relevant filler-gap dependency (Y is the only entity of its type in the image), we treated these examples as fillers. All stimuli were normed with three native speakers and recorded by a fourth native speaker (female). Some actual stimuli corresponding to the images in Figure 1 appear below; relative clauses are in brackets.

- (19)a. Ko fe e tama fi^hfine [ne mamali mogonei]? *Intransitive subject gap*
where DET girl NFT smile now
'Where is the girl who is smiling now?'
- b. Ko fe e puti [ne epoepo he kuli^h]? *Transitive object gap*
where DET cat NFT lick ERG dog
'Where is the cat that the dog is licking?'

- c. Ko fe e kulī [ne epoepo e puti]? *Transitive subject gap*
where DET dog NFT lick ABS cat
'Where is the dog that is licking the cat?'

The experiment was designed and run using the Experigen software (Pillot et al. 2012) on a Macintosh computer in and around Auckland, New Zealand in January, 2014. Each subject saw the stimuli in a random order, and each image was randomly selected to appear either with the canonical order of participants X_1 -Y- X_2 , or the variant flipped about the vertical axis (X_2 -Y- X_1 order). Prior to the experimental portion of the task, subjects were given detailed instructions, both visual (on the computer screen) and auditory, and had the option to replay these instructions indefinitely before moving on. Subjects were then provided with a practice portion of the task, consisting of images and stimuli identical to those that appeared in the experimental portion. They could move on only once they had correctly answered the trial question.

In the experimental portion of the task, the relevant image appeared on screen, followed five seconds later by the auditory stimuli asking the user to identify the relevant entity. Directly beneath each entity was a button that subjects could click to indicate their choice. Reaction times were measured by computing the time elapsed from the onset of the recording to the button click on the screen. Speakers pressed the space bar to progress through the experiment. A brief demographic survey was completed at the end of the survey.

3.2. Results

Before presenting the results, it is important to note that direct comparison of accuracy percentages and response times to similar experiments in better studied languages is difficult. The vast majority of experiments reported in the processing literature have been conducted in the university setting, with researchers testing undergraduates. As has been repeatedly shown (see especially Gagliardi 2012; Clemens et al. 2015), reaction times (RTs) and accuracy rates are notably different in populations not familiar with the fundamentals of experimentation or testing (familiarity with computers, experience with test-taking, etc.). To compound matters, Niuean is not taught in schools or used for any testing purposes whatsoever; as an indication of that, many subjects who took our study were surprised to see the language being used outside a strictly colloquial or liturgical context. With this disclaimer in mind, let us examine the results.

Data were trimmed according to reaction time, with responses taking longer than 30,000ms not included. The responses were further separated according to the type of extraction being performed in each instance: extraction of an absolute subject (S EXT), extraction of an ergative subject (ERG EXT), and extraction of an absolute object (ABS EXT). For each extraction type, we measured the proportion of trials in that extraction type for which the participant answered correctly, and we recorded

the response times associated with each answer. Because there was no significant difference between all response times recorded in the experiment and response times for only those trials in which participants responded accurately (Wilcoxon test, p -value = 0.759), we will only report mean response times for trials in which participants selected the correct response. For S EXT, participants selected the correct response in 72.1% of trials. The mean response time for these trials was 3481ms. For ERG EXT, participants responded accurately to 71.8% of trials. The mean response time for these correct responses was 4377ms. For ABS EXT, participants correctly answered 79.5% of trials. The mean response time for these correct responses was 4419ms. There is no significant difference between the rate of accuracy in participants' response to S EXT, ERG EXT, and ABS EXT, although we observe a trend that participants are more accurate in ABS EXT trials (Fisher's Exact test, $p=0.08$). The mean response time for correctly answered S EXT trials (3481ms) is significantly lower than the mean response times for correctly answered ERG EXT (4377ms) and ABS EXT (4419ms) trials (Wilcoxon test, $p<0.01$). We find no significant difference between ERG EXT and ABS EXT response times, and there is no noticeable trend (Wilcoxon test, $p=0.96$).

(20) Results summary

Condition	Response Time	Accuracy
S EXT	3481ms	72.1%
ERG EXT	4377ms	71.8%
ABS EXT	4419ms	79.5%

4. Discussion of results

Consider first the main question we sought to answer with the experiment: do A' dependencies involving ergative gaps incur a higher processing load than corresponding dependencies involving absolutive object gaps? In Niuean at least, the answer is definitively no. Both types of gaps in transitive clauses (ergative and absolutive object) took longer to process than absolutive subject gaps in intransitive clause, but crucially, there was no significant difference in response time among absolutive object gaps and ergative subject gaps. Returning to the central issue taken up in this paper, this result calls into question the tenability of the processing account to syntactic ergativity. Indeed the relative ease with which ergative gaps are processed in Niuean makes it unlikely that Tongan and Niuean differ merely in their tolerance for complex structures.

Given that it provides a clear-cut case of closely related languages differing in their tolerance for complex structure, crossed-dependency formation in Scandinavian is once again a relevant comparison. Recall that the differential behavior of Swedish and Norwegian with respect to A' -movement from embedded questions exactly mirrors the differential behavior of ergative arguments under A' -extraction in Tongan and Niuean: in Swedish but not Norwegian, all but the lowest extraction

site involved in crossed-dependency formation must be occupied by resumptive pronouns. Crossed-dependencies of the sort on display here have been independently argued theoretically and experimentally to be both complex in a formal-language-theoretic sense (Miller 1991) and difficult to process (Fodor 1978; Engdahl 1979). Additionally, there exist independent arguments that the relevant resumptive pronouns in Swedish are essentially devices to facilitate processing (see Engdahl 1982; Asudeh 2004), so that it follows that Norwegian and Swedish simply have differential tolerance for complex structure. The relevance of these data are that they show that processing-based accounts can be successful in explaining data that are similar to what we have observed in Tongan and Niuean. This, in turn, suggests that our failure to observe increased processing costs for the extraction of ergative arguments is truly because processing is not at stake here: in those cases where processing factors are actually implicated in differential extraction patterns, they are readily and immediately observable.

The results from our Niuean study are congruent with related results from Avar (Polinsky et al. 2012, Polinsky to appear b), a Nakh-Dagestanian language that also exhibits morphological but not syntactic ergativity. Unlike Niuean, Avar does not have a closely related language with syntactic ergativity that could be used for comparison, although the similarity of the results to those obtained in Niuean suggest we are dealing with pervasive cross-linguistic trends. Polinsky et al. (2012), Polinsky to appear b conducted both a self-paced reading study and a picture-matching study of the processing of relative clauses in Avar. The self-paced reading study revealed that Avar participants processed relative clauses with absolutive-subject gaps faster than the corresponding relative clauses with absolutive object gaps or ergative gaps—just as in Niuean. At the same time, the ergative and absolutive object gaps were processed at roughly the same reading speed, albeit with a different time course. The results from the picture-matching study were similar, with a clear preference, measured in terms of response time, for absolutive subject gaps. Ergative and absolutive object gaps showed no significant differences in terms of response time or accuracy.

Beyond their relevance for arguing against processing-based accounts, the results from these studies, especially when taken in conjunction, have ramifications for another well known issue in the processing literature. Specifically, it has been repeatedly noted that there exists a so-called *subject processing advantage*, which refers to the general, strong processing advantage, measured variously by reaction time, neuroimaging, etc., associated with A'-dependencies that involve a subject gap over those that involve an object gap. Such effects have been experimentally observed in at least English (King and Kutas 1995; Traxler et al. 2002 a.o.), German (Schwartz 2007, and further references therein), Dutch (Frazier 1987, 1989), Japanese (Miyamoto and Nakamura 2003; Ishizuka et al. 2003), Korean (Kwon et al. 2006, 2010, 2013), Russian (2011; Levy et al. 2013 and further references therein), and Turkish (Demiral et al. 2008; Özge et al. 2009). All of these languages exhibit nominative-accusative alignment, raising the question of whether the SPA is truly about subjecthood or rather has to do with morphological cueing effects associated with the alignment of case morphology in these languages. Specifically, the appear-

ance of an accusative marked argument, an object, is contingent on the appearance of a nominative marked argument, a subject, so that accusative case is sometimes referred to as *dependent* case (Marantz 1984). This dependence relation between nominative and accusative case has been shown to provide a *cue* to the processor in the parsing of certain filler-gap dependencies. The logic is as follows. The presence of an accusative argument in a given clause cues the processor that the clause is transitive and that a nominative argument should be projected. Accordingly, nominative gaps in, for example, relative clauses are predicted to be easier to parse because the presence of accusative case cues the parser to the presence of a nominative argument that, if absent, must be the head of the relative clause. Several studies provide empirical evidence in favor of this cueing effect in the processing of relative clauses, most notably in Korean (Kwon et al. 2006) and Japanese (Ueno and Garnsey 2008).⁷ Because this morphological cueing effect intervenes to facilitate the processing of nominative gaps, it is difficult to tease apart the relative import of morphological case and other *subject*-properties on the SPA. In other words, it is unclear whether the SPA exists independent of morphological cueing.

In this context, our results from the morphologically ergative languages Niuean and Avar have the potential to differentiate the relative import of morphological cueing and other subjecthood properties on the SPA. In these languages, it is the case borne by transitive subjects, ergative, that is dependent: it is only in those (transitive) clauses with an absolutive object that ergative case surfaces on the subject. Consequently, the case-cueing effect, if it exists in these languages, should be expected to be triggered by ergative case: the presence of an ergative argument informs the parser that there must be an absolutive argument projected in the clause. In contrast, if the SPA exists independent of case cueing, it should work in these languages to facilitate the processing of both ergative-subject gaps and absolutive-subject gaps, but not absolutive-object gaps. Case cueing and the SPA are thus at odds in these languages.

The results from Niuean and Avar seem to confirm, or at least are consistent with, the conclusion that morphological cueing and the SPA are two independent phenomena. Specifically, we observed that absolutive object gaps and ergative gaps, where the SPA (if it exists independent of morphological cueing) and morphological cueing are at odds, are more difficult to process than absolutive subject gaps, where the SPA is expected to advantage processing. That is, in those cases where the two effects are predicted to be at odds, we observed correspondingly slower processing than in the case where we would expect the SPA to advantage processing.⁸ Further

⁷See also Bornkessel-Schlesewsky and Schlewsky 2009 and Skopeteas et al. 2012 for evidence that morphological case is a strong cue in processing.

⁸The transitivity effect observed in Niuean and Avar, which held even though the intransitive conditions were of the same length as the transitive ones, with a PP in place of the object, has been replicated in studies on accusative languages as well (see, for example, ?, who document a transitivity penalty in Russian). This suggests that morphological cueing and the SPA may not be the only factors at play in these data, although more work is certainly needed.

support for this conclusion comes from results on processing of relative clauses in Mayan languages (Norcliffe 2009; Clemens et al. 2015). In the languages studied here, ergativity is encoded only on predicate heads, not via case marking, so that there is a corresponding lack of morphological cueing. Clemens et al. (2015) show that in the absence of such cueing effects, the SPA leads to a clear advantage for the processing of absolutive subject gaps and ergative gaps over absolutive object gaps.

5. Conclusion

The explicit aim of this investigation was the determination of whether syntactic ergativity might yield to a processing-based account along the lines suggested by Hawkins (2014) for related phenomena such as accessibility of different constituents to relative clause or wh-question formation. The experimental evidence we gathered on the processing of ergative and absolutive gaps in Niuean casts doubt on a processing-based explanation of ergativity, especially when taken in conjunction with similar results in Avar (Polinsky et al. 2012) as well as Chol and Q'anjob'al (Clemens et al. 2015). In light of this conclusion, it is especially important to revisit the various syntactic approaches to syntactic ergativity, and in this domain too Niuean and Tongan present an excellent test case. It should be possible to isolate the source(s) of syntactic ergativity, whatever it may be, by comparing these two exceptionally similar languages.

Of course, restrictions on A'-extraction and the notion of subject constitute only a small fraction of the topics that experimental investigation into ergative languages has the potential to inform our understanding of. To list a few examples, our understanding of the distribution of ergativity, including a number of typological gaps (accusative languages with syntactic ergativity, languages with an accusative case system but an ergative agreement pattern, etc.), of the nature of ergative case (structural vs. inherent), of the heterogeneity of ergative languages (absolutive sometimes appears as a default case and sometimes as an analog of nominative) could all stand to benefit from further experimental investigation (see Longenbaugh and Polinsky to appear for more discussion). While these areas all ostensibly concern properties of ergative languages, as we saw in the study presented here, the results can have ramifications that reach far beyond the original questions. The work presented and discussed herein ought thus be viewed as a preliminary demonstration of the importance of the experimental study of ergative languages.

References

- Aldridge, Edith. 2008. Generative approaches to ergativity. *Language and Linguistics Compass* 2:966–995.
- Aldridge, Edith Catherine. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.
- Anderson, Stephen R. 1976. On the notion of subject in ergative languages. In *Subject and topic*, ed. Charles N. Li, volume 1, 23. New York: Academic

- Press.
- Asudeh, Ash. 2004. Resumption as resource management. Doctoral Dissertation, Stanford University.
- Bhatt, Rajesh. 2002. The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics* 10:43–90.
- Bittner, Maria. 1994. *Case, scope, and binding*. Springer.
- Bittner, Maria, and Ken Hale. 1996. Ergativity: Toward a theory of a heterogeneous class. *Linguistic Inquiry* 27:531–604.
- Bobaljik, Jonathan David, and Phil Branigan. 2006. Eccentric agreement and multiple case checking. In *Ergativity: emerging issues*, ed. Alana Johns, Diane Massam, and Juvenal Ndayiragije, 47–77. Springer.
- Bornkessel-Schlesewsky, Ina, and Matthias Schlesewsky. 2009. *Processing syntax and morphology: A neurocognitive perspective*. Oxford University Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken hale: A life in language*, ed. Michael Kenstowicz, 1–52. MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. MIT Press.
- Clemens, Lauren Eby, Jessica Coon, P Mateo Pedro, AM Morgan, Maria Polinsky, Gabrielle Tandet, and Matt Wagers. 2015. Ergativity and the complexity of extraction: A view from Mayan. *Natural Language & Linguistic Theory* 33:417–467.
- Comrie, Bernard, and Edward L Keenan. 1979. Noun phrase accessibility revisited. *Language* 55:649–664.
- Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. 2015. The role of case in A-bar extraction asymmetries: evidence from Mayan. *Linguistic Variation* 14:179–242.
- Coon, Jessica, and Omer Preminger. 2012. Towards a unified account of person splits. In *Proceedings of the 29th West Coast Conference on Formal Linguistics (WCCFL 29)*, 310–318. Cascadilla Press.
- Creider, Chet. 1986. Constituent-gap dependencies in norwegian: An acceptability study. In *Diversity and diachrony*, ed. David Sankoff, 415–24. John Benjamin.
- Demiral, Şükrü Barış, Matthias Schlesewsky, and Ina Bornkessel-Schlesewsky. 2008. On the universality of language comprehension strategies: Evidence from Turkish. *Cognition* 106:484–500.
- Engdahl, Elisabet. 1979. The nested dependency constraint as a parsing principle. In *Papers presented to emmon bach by his students*, ed. Elisabet Engdahl and Mark J. Stein, 76–87. University of Massachusetts, Amherst.
- Engdahl, Elisabet. 1982. Restrictions on unbounded dependencies in Swedish. In *Readings on unbounded dependencies in Scandinavian languages*, 151–174. Universitetet i Umeå.
- Fodor, Janet Dean. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9:427–473.

- Frazier, Lyn. 1987. Syntactic processing: evidence from Dutch. *Natural Language & Linguistic Theory* 5:519–559.
- Frazier, Lyn, and Charles Clifton Jr. 1989. Successive cyclicity in the grammar and the parser. *Language and cognitive processes* 4:93–126.
- Gagliardi, Ann C. 2012. Input and intake in language Acquisition. Doctoral Dissertation, University of Maryland.
- Gerds, Donna. 1988. *Object and absolute in Halkomelem Salish*. Garland.
- Gordon, Peter C, and Randall Hendrick. 2005. Relativization, ergativity, and corpus frequency. *Linguistic Inquiry* 36:456–463.
- Haji-Abdolhosseini, Mohammad, Diane Massam, and Kenji Oda. 2002. Number and events: Verbal reduplication in Niuean. *Oceanic Linguistics* 41:475–492.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford University Press.
- Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford University Press.
- Hewitt, George B. 1979. *Abkhaz*. Elsevier.
- Hulsey, Sarah, and Uli Sauerland. 2006. Sorting out relative clauses. *Natural Language Semantics* 14:111–137.
- Ishizuka, Tomoko, Kentaro Nakatani, and Edward Gibson. 2003. Relative clause extraction complexity in Japanese. Poster Presented at the 16 the Annual CUNY Conference on Human Sentence Processing.
- Keenan, Edward L, and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic inquiry* 8:63–99.
- King, Jonathan, and Marta Kutas. 1995. Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience* 7:376–395.
- Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. In *The limits of syntax*, ed. Peter W. Culicover and Louise McNally, 241–280. Academic Press.
- Kluender, Robert. 2004. Are subject islands subject to a processing account. In *Proceedings of the West Coast Conference on Formal Linguistics*, volume 23, 475–499.
- Kwon, Nayoung, Peter C Gordon, Yoonhyoung Lee, Robert Kluender, and Maria Polinsky. 2010. Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of prenominal relative clauses in Korean. *Language* 86:546–582.
- Kwon, Nayoung, Robert Kluender, Marta Kutas, and Maria Polinsky. 2013. Subject/object processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language* 89:537–585.
- Kwon, Nayoung, Maria Polinsky, and Robert Kluender. 2006. Subject preference in Korean. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, ed. D. Montero D. Baumer and M. Scanlon, 1–14. Cascadilla Press.
- Larson, Bradley, Nicholas Longenbaugh, and Maria Polinsky. 2015. Niuean sub-

- ject/object parity and the labelling algorithm. In *Wccfl talk*.
- Legate, Julie Anne. 2006. Split absolutive. In *Ergativity: emerging issues*, ed. Alana Johns, Diane Massam, and Juvenal Ndayiragije, 143–171. Springer.
- Legate, Julie Anne. 2008. Morphological and abstract case. *Linguistic Inquiry* 39:55–101.
- Levy, Roger, Evelina Fedorenko, and Edward Gibson. 2013. The syntactic complexity of russian relative clauses. *Journal of memory and language* 69:461–495.
- Longenbaugh, Nicholas, and Maria Polinsky. 2015. Equidistance returns: on the parity of subject and object in Niuean. *Unpublished Manuscript*.
- Longenbaugh, Nicholas, and Maria Polinsky. to appear. Experimental approaches to ergative languages. In *The handbook of ergativity*, ed. Jessica Coon, Diane Massam, and Lisa Travis.
- Maling, Joan, and Annie Zaenen. 1982. A phrase structure account of Scandinavian extraction phenomena. In *The nature of syntactic representation*, ed. Pauline Jacobson and Geoffrey Pullum, 229–282. Springer.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. MIT Press.
- Markman, Vita G, and Pavel Grashchenkov. 2012. On the adpositional nature of ergative subjects. *Lingua* 122:257–266.
- Massam, Diane. 2000. VSO and VOS: Aspects of Niuean word order. In *The syntax of verb initial languages*, ed. Andrew Carnie and Eithne Guilfoyle, 97–116. Oxford University Press.
- Massam, Diane. 2001. On predication and the status of subjects in Niuean. In *Objects and other subjects*, 225–246. Springer.
- Massam, Diane, Donna Starks, and Ofania Ikiua. 2006. On the edge of grammar: discourse particles in Niuean. *Oceanic linguistics* 45:191–205.
- Miller, Philip H. 1991. Scandinavian extraction phenomena revisited: Weak and strong generative capacity. *Linguistics and Philosophy* 14:101–113.
- Miyamoto, Edson T, and Michiko Nakamura. 2003. Subject/object asymmetries in the processing of relative clauses in Japanese. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, ed. Gina Garding and Mimu Tsujimura, 342–355. Cascadilla Press.
- Norcliffe, Elisabeth. 2009. Head-marking in usage and grammar: a study of variation and change in Yucatec Maya. Doctoral Dissertation, Stanford University.
- Otsuka, Yuko. 2000. Ergativity in Tongan. Doctoral Dissertation, University of Oxford.
- Özge, Duygu, Theodoros Marinis, and Deniz Zeyrek. 2009. Comprehension of subject and object relative clauses in monolingual Turkish children. In *Essays on Turkish Linguistics, Proceedings of the 14th International Conference on Turkish Linguistics*, ed. Sila Ay, Özgür Aydın, İclâl Ergenç, Seda Gökmen, Selçuk İşsever, and Dilek Peçenek. Harrasowitz Verlag.
- Pesetsky, David Michael. 1982. Paths and categories. Doctoral Dissertation, MIT.
- Pillot, Carl, Greg Scontras, and Lauren Clemens. 2012. ExperigenRT. <http://perma.cc/67VV-NMJL>.
- Polinsky, Maria. 2011. Reanalysis in adult heritage language: A case for attrition.

Studies in Second Language Acquisition 33:305–328.

- Polinsky, Maria. to appear a. Syntactic ergativity. In *The Blackwell Companion to Syntax*. Blackwell.
- Polinsky, Maria. to appear b. *Deconstructing ergativity*. Oxford University Press.
- Polinsky, Maria, Carlos Gomez Gallo, Peter Graff, and Ekaterina Kravtchenko. 2012. Subject preference and ergativity. *Lingua* 122:267–277.
- Potsdam, Eric, and Maria Polinsky. 2011. Questions and word order in Polynesian. In *Topics in Oceanic morphosyntax*, ed. Claire Moyses-Faurie and Joachim Sabel, 107–134. De Gruyter Mouton.
- Richards, Norvin. 1997. What moves where when in which language? Doctoral Dissertation, MIT.
- Schwartz, Florian. 2007. Processing presupposed content. *Journal of Semantics* 24:373–416.
- Seiter, William J. 1980. Issues in Niuean Syntax. Doctoral Dissertation, University of California, San Diego.
- Skopeteas, Stavros, Gisbert Fanselow, and Rusudan Asatiani. 2012. Case inversion in Georgian: Syntactic properties and sentence processing. In *Case, Word Order and Prominence*, ed. Monique Lamers and Peter de Swart, 145–171. Springer.
- Sperlich, Wolfgang. 1995. Is Niuean an endangered language species? *Journal of Educational Studies* 17:37–55.
- Sperlich, Wolfgang. 2005. Will cyberforms save endangered languages? A Niuean case study. *International Journal of the Sociology of Language* 172:51–77.
- Stepanov, Arthur. 2004. Ergativity, case and the minimal link condition. In *Minimality effects in syntax*, ed. Arthur Stepanov, Gisbert Fanselow, and Ralf Vogel, 367–399. Walter de Gruyter.
- Testeleets, Yakov, Peter Arkadev, Alexander Letuchij, and Nina Sumbatova, ed. 2009. *Aspekty polisintetizma: Očerki po grammatike adygejskogo jazyka*. RGGU.
- Traxler, Matthew J, Robin K Morris, and Rachel E Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47:69–90.
- Ueno, Mieko, and Susan M Garnsey. 2008. An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes* 23:646–688.
- Wiltschko, Martina. 2006. On 'ergativity' in Halkomelem Salish. In *Ergativity: emerging issues*, ed. Alana Johns, Diane Massam, and Juvenal Ndayiragije, 197–228. Springer.
- Zawiszewski, Adam, Eva Gutiérrez, Beatriz Fernández, and Itziar Laka. 2011. Language distance and non-native syntactic processing: Evidence from event-related potentials. *Bilingualism: Language and Cognition* 14:400–411.

EXTRAPOSITION IN MALAGASY*

Eric Potsdam
University of Florida
potsdam@ufl.edu

Daniel Edmiston
University of Chicago
danedmiston@uchicago.edu

This paper documents some of the basic properties of extraposition (right peripheral positioning) in Malagasy. Extraposition is generally optional for a wide range of constituents; however, it is obligatory for clausal constituents. We offer a prosodic explanation for this obligatoriness: leaving the clause in its base position would yield an ill-formed prosodic structure. We also argue that extraposition in Malagasy is derived by A' movement. Various reconstruction effects support this claim.

1. Introduction

We use EXTRAPOSITION (EX) as a descriptive term to refer to the non-canonical placement of certain constituents (EXPs) in a right-peripheral position. There are a number of phenomena that are called extraposition in the literature and it is not clear that they all have the same analysis (Baltin 2006, Sheehan 2010). (1a) illustrates SIMPLE EXTRAPOSITION, in which predicate-related material appears in a right peripheral position. (1b) illustrates EXTRAPOSITION FROM NP and (1c) illustrates HEAVY XP SHIFT. We will be focusing on simple extraposition.

- (1) a. He said yesterday [that they were not prepared].
b. A man walked in [who was wearing a red hat].
c. Bob put on the table [all the gifts that his wife insists that they buy].

This paper investigates extraposition in Malagasy, a predicate-initial language spoken by approximately 14 million people on the island of Madagascar. Malagasy has simple extraposition, which one can easily see in the word order. Canonical word order is VOXS, (2), with various elements appearing in the predicate-internal position between the object and the subject. These predicate-internal elements can optionally extrapose, yielding VOSX order, (3). Here and below, the constituents of interest are bold-faced.

* We thank our Malagasy consultants Bodo and Voara Randrianasolo and the audience at AFLA 22. Examples come from our own fieldwork unless otherwise indicated.

- (2) a. Nametraka voninkazo **teo ambonin' ny latabatra** i Koto
 PAST.put flower LOC on DET table Koto
 'Koto put flowers on the table.'
- b. Mananatra ny zafikeliny **mandrakariva** i Dadabe
 admonish DET grandchild.3SG always grandpa
 'Grandpa admonishes his grandchildren constantly.'
- (3) a. Nametraka voninkazo i Koto **teo ambonin' ny latabatra**
 PAST.put flower Koto LOC on DET table
 'Koto put flowers on the table.'
- b. Mananatra ny zafikeliny i Dadabe **mandrakariva**
 admonish DET grandchild.3SG grandpa always
 'Grandpa admonishes his grandchildren constantly.'

CPs in Malagasy have a special status with respect to extraposition in that they obligatorily extrapose (Keenan 1976, others), (4).

- (4) a. Manantena Rabe **fa hividy fiara aho**
 hope Rabe that buy car 1SG.NOM
 'Rabe hopes that I will buy a car.'
- b. *Manantena **fa hividy fiara aho** Rabe
 hope that buy car 1SG.NOM Rabe

The goals and content of this paper are as follows. Section 2 documents the empirical patterns of Malagasy extraposition, both the syntactic distribution and semantic consequences. These have not been previously described. Section 3 considers the analysis of EX in Malagasy and argues in favor of a movement approach. EXPs are moved from a predicate-internal position to their surface position and are not base-generated there. Section 4 more carefully considers the obligatoriness of CP extraposition and offers a prosodic explanation. Section 5 restates our findings and mentions areas for future investigation.

2. Extraposition

The syntactic distribution of extraposed phrases in Malagasy can be summarized as follows: EX is impossible for objects, obligatory for full clauses, and optional for other elements, including various complements, adverbials, and controlled clauses. We illustrate these observations in the data below.

(5) shows that EX is impossible for direct objects.

- (5) Namono (**ny akoho**) Rasoana (***ny akoho**)
 PAST.kill DET chicken Rasoana DET chicken
 'Rasoana killed the chicken.'

As we have already seen in (4), EX is obligatory for full CPs (Keenan 1976, Pearson 2001, Law 2007, Potsdam and Polinsky 2007). This is true for both complement clauses in (6) (repeated from (4)) and adverbial clauses, (7).

- (6) a. Manantena Rabe **fa hividy fiara aho**
 hope Rabe that buy car 1SG.NOM
 ‘Rabe hopes that I will buy a car.’
 b. *Manantena **fa hividy fiara aho** Rabe
 hope that buy car 1SG.NOM Rabe
- (7) a. Tsy nianatra Rabe **satria narary ny vadiny**
 NEG study Rabe because sick DET spouse.3SG
 ‘Rabe didn’t study because his wife was sick.’
 b. *Tsy nianatra **satria narary ny vadiny** Rabe
 NEG study because sick DET spouse.3SG Rabe

EX is optional for other dependents in the predicate (Rajaonarimanana 1995:87). This includes various kinds of PPs (locatives, recipients, goals, instruments, benefactives, material themes), adverbials (temporal, manner, locative, reason, and frequency), standards of comparison, and controlled clauses (complements and adjuncts). Representative data are given below (we do not illustrate every possibility for lack of space). (8) shows extraposition of a locative PP and a recipient PP. (9) illustrates extraposition of a manner adverbial and a frequency adverb. (10) illustrates extraposition of a standard of comparison. Extraposition of controlled clauses is seen later, in (33).

- (8) a. Nitoetra (**tamin’ ity trano ity**) nandritran’ ny
 PAST.live PREP DEM house DEM during DET
 raopolo taona ry Ratsimba (**tamin’ ity trano ity**)
 twenty year DET Ratsimba PREP DEM house DEM
 ‘The Ratsimbas lived in this house for twenty years.’
 b. Nanome vola (**ho an-dRabe**) aho (**ho an-dRabe**)
 PAST.give money PREP PREP-Rabe 1SG.NOM PREP PREP-Rabe
 ‘I gave money to Rabe.’
- (9) a. Namono akoho (**tamin-katezerana**) ny zaza
 kill.PAST chicken PREP-anger DET child
 (**tamin-katezerana**)
 PREP-anger
 ‘The child killed the chickens angrily.’
 b. Tsy mandamina ny trano (**matetika**) Rakoto (**matetika**)
 NEG arrange DET house often Rakoto often
 ‘Rakoto often does not put the house in order.’

- (10) Hendry (**noho ny zandriny**) i Koto (**noho ny zandriny**)
 wise than DET sibling.3SG Koto than DET sibling.3SG
 ‘Koto is wiser than his younger sibling.’

As has been noted in passing by a handful of researchers, EX has semantic consequences. Extraposed constituents are backgrounded/presupposed (Paul and Rabaovololona 1998, Pearson 2001, Kalin 2009). The only explicit evidence for this claim in the literature comes from Pearson 2001:180:

- (11) a. Namaky boky **tany an-tokotany** ve i Tenda?
 read book LOC PREP-garden Q Tenda
 ‘Was reading a book in the garden what Tenda was doing?’
 b. Namaky boky ve i Tenda **tany an-tokotany**
 read book Q Tenda LOC PREP-garden
 ‘Was reading a book what Tenda was doing in the garden?’

Both (11a) and (11b) translate broadly as ‘Was Tenda reading a book in the garden?’. In (11a), the PP ‘in the garden’ is inside the predicate and the question more accurately asks, ‘Was reading a book in the garden what Tenda was doing?’ In (11b), the PP is extraposed. As a result, it is backgrounded and the question is best paraphrased as ‘Was reading a book what Tenda was doing in the garden?’. That is, it is presupposed that Tenda was doing something in the garden and the question asks whether what he was doing there is reading a book.

This characterization of EX correctly predicts that certain elements cannot extrapose because they cannot be backgrounded. This includes wh-phrases and answers to questions. (12) shows that wh-phrases cannot occur in extraposed position. This follows if they are not presuppositional (Fitzpatrick 2005).

- (12) a. Lasa nody (**oviana**) Rabe (***oviana**)?
 gone PAST.go.home when Rabe when
 ‘Rabe went home when?’
 b. Nataon’ i Jehovah (**tamin’ iza**) ilay fifanekena (***tamin’ iza**)?
 do.PASS Jehovah PREP who DEM covenant PREP who
 ‘Jehovah made that covenant with who?’

The question/answer pair in (13) shows that the answer to a wh-question also cannot be extraposed, A2. The new information must appear unextraposed, A1.

- (13) Q: Oviana no lasa nody Rabe?
 when FOC gone PAST.go.home Rabe
 ‘When did Rabe go home?’

- A1: Lasa nody **omaly hariva** Rabe
gone PAST.go.home yesterday evening Rabe
'Rabe went home yesterday evening.'
- A2: #Lasa nody Rabe **omaly hariva**
gone PAST.go.home Rabe yesterday evening

We note in passing that extraposed CPs need not be backgrounded and can function as an answer to a question, (14). This is unsurprising since the word order in A1 without extraposition is simply ungrammatical.

- (14) Q: Inona no notenenan-dRabe momba ahy?
what FOC PAST.say.PASS-Rabe about 1SG.ACC
'What did Rabe say about me?'
- A1: *Niteny **fa miasa tsara ianao** Rabe
said that work well 2SG.NOM Rabe
'Rabe said that you work hard.'
- A2: Niteny Rabe **fa miasa tsara ianao**
said Rabe that work well 2SG.NOM

To summarize, extraposed constituents appear clause-finally, after the subject in Malagasy. Extraposition is impossible for objects, obligatory for full clauses, and optional otherwise. Except when extraposition is obligatory, extraposed constituents are backgrounded.

3. Analysis

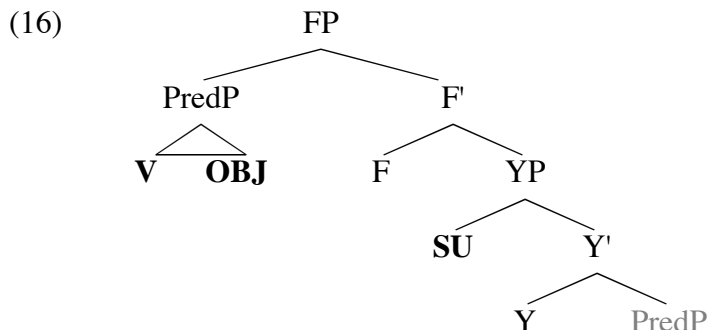
There is no shortage of analytical approaches to EX. We consider two here: an A' movement analysis (Ross 1967, Baltin 1982, Büring and Hartmann 1997, others), and a base-generation analysis (Rochemont and Culicover 1990).¹

Before developing instantiations of these analyses, we introduce our assumptions about Malagasy clause structure, in (15). We adopt a predicate-fronting analysis of VOS word order (Massam and Smallwood 1997, Rackowski and Travis 2000, Pearson 2001, others).

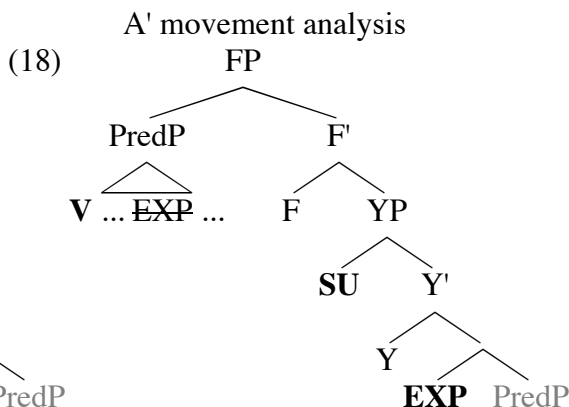
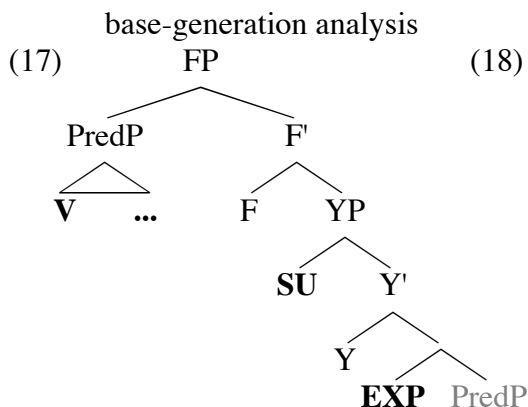
- (15) a. underlying SVO order
b. subject occupies the specifier of a phrase YP above the predicate
c. VOS is derived by leftward movement of a predicate constituent (PredP) to a position above the subject
d. PredP reconstructs to its base position at LF (Massam 2000, Potsdam 2007)

¹ Three more recent analyses that we will not explore given space considerations are a stranding analysis (Kayne 1994, Wilder 1996, Sheehan 2010), an ellipsis analysis (de Vries 2009), and a PF movement analysis (Göbbel 2007).

VOS clauses begin as SVO with the subject occupying a position that we call spec,YP for concreteness. VOS is derived by leftward movement of a predicate constituent, PredP. This movement is A' movement and PredP reconstructs to its base position at LF (Massam 2000, Potsdam 2007). A derivation for VOS is shown in (16).



Under a base-generation analysis of EX, the extraposed constituent is base-generated outside the predicate in a position to the right of the subject, (17). Under a movement analysis, the extraposed constituent is base-generated inside the predicate. It A' moves to a position outside the predicate. PredP then fronts, an instance of remnant movement, (18).



3.1. Evidence for A' movement

Evidence for the movement analysis comes from data showing that extraposed constituents behave as though they are in a predicate-internal position. We call these RECONSTRUCTION facts and they are summarized in (19). They are accounted for under a movement analysis on the assumption that EXPs originate in the predicate-internal position and can reconstruct to this position. They are unexplained on the base-generation analysis since EXPs are never in a predicate-internal position.

- (19) a. EXPs show reconstruction for syntactic selection
 b. EXPs can contain a pronoun bound by an object
 c. EXPs require reconstruction for Binding Principles
 d. EXPs require reconstruction for NPI licensing
 e. EXPs are not islands to extraction

The first argument comes from the observation that EXPs show reconstruction for syntactic selection. Complement selection is widely assumed to require locality and some EXPs are clearly complements:

- (20) Tezitra (**amin' ny mpiasa**) i Dada (**amin' ny mpiasa**)
 angry PREP DET worker father PREP DET worker
 'Father is angry with the worker.'

In the same vein, idiom pieces can extrapose and selection is required for idiom formation (O'Grady 1998, Bruening 2010):

- (21) Mampiditra ahy (**amin' ny kizo**) ianao (**amin' ny kizo**)
 insert 1SG.ACC PREP DET alley 2SG PREP DET alley
 lit. "You're inserting me into a dangerous alley."
 'You're trying to trick me.' (Winterton 2011:90)

The second argument comes from variable binding. EXPs can contain a pronoun bound by an object.² In (22), the extraposed PP contains a pronoun bound by the quantified object. In (23), the extraposed CP contains a bound pronoun. In a predicate fronting derivation, the object does not c-command EXP in these examples unless EXP can reconstruct back inside the predicate.

- (22) Nametraka ny zazakely tsirairay_i (**teo ambonin' ny**
 PAST.put DET child each LOC in DET
fandria-ny_i) ny mpitsabo (**teo ambonin' ny fandria-ny_i**)
 bed-3SG DET nurse LOC in DET bed-3SG
 'The nurse put each child_i in his_i bed.'
- (23) Niteny tamin' ny zazalahy tsirairay_i aho
 PAST.say PREP DET boy each 1SG
fa hanoroka azy_i Raso
 that FUT.kiss 3SG Raso
 'I told each boy_i that Raso will kiss him_i.'

² The subject may also bind a pronoun in an extraposed clause (Zribi-Hertz and Mbolatianavalona 1999, Law 2007); however, this is predicted for the structures under consideration regardless of whether or not there is reconstruction.

The third argument comes from the Binding Principles. Predicate-internal pronominal objects trigger Condition C violations with respect to R-expressions contained in an extraposed constituent. This is shown in (24) for an optionally extraposed PP and in (25) for an obligatorily extraposed CP. The object does not c-command into EXP unless it reconstructs.

- (24) a. Nampahatsiahy an-dRabe_i (**momba ny fivoriana-ny_i**)
 remind ACC-Rabe about DET meeting-3SG
 aho (**momba ny fivoriana-ny_i**)
 1SG.NOM about DET meeting-3SG
 ‘I reminded Rabe_i about his_i meeting.’
- b. *Nampahatsiahy azy_i (**momba ny fivoriana-dRabe_i**)
 remind 3SG.ACC about DET meeting-Rabe
 aho (**momba ny fivoriana-dRabe_i**)
 1SG.NOM about DET meeting-Rabe
 ‘*I reminded him_i about Rabe_i’s meeting.’
- (25) a. Nampahatsiahy an-dRabe_i aho **fa efa nahita azy_i Rasoa**
 PAST.remind ACC-Rabe 1SG that PFV PST.see 3SG Rasoa
 ‘I reminded Rabe_i that Rasoa already saw him_i.’
- b. *Nampahatsiahy azy_i aho **fa efa nahita an-dRabe_i Rasoa**
 PAST.remind 3SG 1SG that PFV PST.see ACC-Rabe Rasoa
 ‘*I reminded him_i that Rasoa already saw Rabe_i.’

The fourth argument comes from Negative Polarity Item (NPI) licensing. We assume that NPIs must be c-commanded by negation and that Malagasy negation *tsy* ‘NEG’ is inside the predicate as in (26a) and not external to the predicate as in (26b).

- (26) a. ✓ [[NEG PREDICATE] SUBJECT]
 b. ✗ [NEG [PREDICATE SUBJECT]]

Evidence for (26a) comes from the observations that subject NPIs are not allowed (unless licensed by genericity or modality) (Paul 2005), (27), and coordination shows that negation forms a constituent with the verb, (28).

- (27) *Tsy nanongo an’ i Koto n’iza n’iza
 NEG PAST.pinch ACC Koto anyone
 (‘No one pinched Koto.’) (Paul 2005:363, (13a))
- (28) [[Tsy maty] ary [tsy manampahataperana]] Izy
 NEG dead and NEG ended 3SG.NOM
 ‘He (God) is not dead, he is eternal.’

Nevertheless, predicate-internal NPIs may extrapose, (29). Reconstruction of the extraposed NPI is required on the assumption that the surface position of EXPs is outside the predicate and not c-commanded by negation.

- (29) Tsy nandroso vary (**tamin' n'iza n'iza**) i Be (**tamin' n'iza n'iza**)
 NEG PAST.serve rice PREP anyone Be PREP anyone
 'Be didn't serve rice to anyone.'

The final argument comes from extraction. A widely-cited generalization about Malagasy wh-questions is that only subjects and adjuncts can be questioned (Keenan 1976, 1995, others):

- (30) a. Iza no hividy fiara?
 who FOC FUT.buy car
 'Who will buy a car?'
 b. Rahoviana no hividy fiara Rabe?
 when.FUT FOC FUT.buy car Rabe
 'When will Rabe buy a car?'
 c. *Inona no hividy Rabe?
 what FOC FUT.buy Rabe
 ('What will Rabe buy?')

(31) shows that an adjunct wh-phrase can be long-distance questioned from inside an extraposed CP. Given the tense morphology on the wh-phrase and the two verbs, this must be a question about the time of the complement clause event; the wh-phrase must be originating in the embedded clause.

- (31) Rahoviana_i no mihevitra Rabe **fa hividy fiara t_i i Soa?**
 when.FUT FOC PRES.think Rabe that FUT.buy car Soa
 'When does Rabe think that Soa will buy a car?'
 (must be a question about the time of car-buying)

Extraction from EXP is problematic for a base-generation analysis of EX because movement from EXP in its extraposed position would violate Huang's (1982) Condition on Extraction Domains (CED), (32), which prohibits extraction from a non-complement. Under the A' movement analysis, movement is taking place from the extraposed clause when it is in its predicate-internal position. It subsequently extraposes.

- (32) *Condition on Extraction Domain* (CED) (Huang 1982:505)
 A phrase may be extracted out of a domain B only if B is properly governed

To summarize this section, reconstruction effects indicate that extraposed elements behave syntactically as though they were in a predicate-internal position. Such facts are incompatible with a base-generation analysis in which EXPs originate in a predicate-external position and require one in which EXPs are in a predicate-internal position at some point in the derivation.

4. Obligatoriness of CP Extraposition

We have argued thus far that, with respect to reconstruction facts, extraposition of CPs (CPEX) and extraposition of other phrases (EX) behave alike. That is, they could be a single phenomenon. There are two differences between CPEX and EX which prevent adopting this view of things. First, EX is optional but CPEX is obligatory and, second, EX backgrounds the EXP but CPEX does not. In this section, we address the first difference, arguing that it has an independent explanation.

We start with some additional data. As we have seen, CPEX is obligatory for both complement and adjunct CPs, (6) and (7). What has not been explicitly noticed before is that CPEX is only obligatory when the CP contains an overt clause-final subject. In situations where the CP lacks a clause-final subject, extraposition again becomes optional. We illustrate with three such situations. (33) demonstrates that controlled clauses, both complements and adjuncts, only optionally extrapose. In (34), subject-less existential clauses also need not extrapose. Finally, (35) illustrates a construction which Potsdam and Polinsky 2007 analyzes as topic drop in finite clauses. Lacking an overt subject, these clauses too do not have to extrapose.

- (33) a. Manantena (**hianatra teny anglisy**) Rabe
 PRES.hope FUT.learn language English Rabe
 (**hianatra teny anglisy**)
 FUT.learn language English
 ‘Rabe hopes to learn English.’
- b. Mianatra mafy (**mba hahazo karama be**)
 study hard COMP.IRR get wages big
 ilay mpianatra (**mba hahazo karama be**)
 DEM student COMP.IRR get wages big
 ‘The student studies hard in order to earn a big salary.’ (Paul 2000:94)
- (34) Milaza (**fa misy gidro any an-tsena**) Rabe
 PRES.say that exist lemur LOC PREP-market Rabe
 (**fa misy gidro any an-tsena**)
 that exist lemur LOC PREP-market
 ‘Rabe says that there are lemurs at the market.’

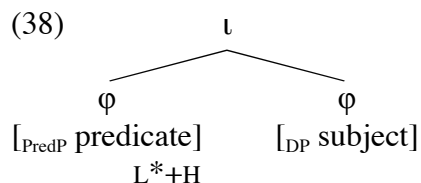
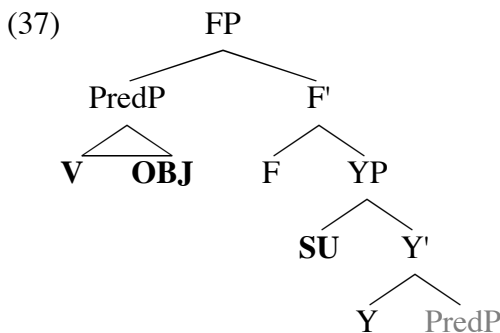
- (35) a. Milaza (fa nahita gidro tany an-tsena) Rabe
 PRES.say that PAST.saw lemur LOC PREP-market Rabe
 (fa nahita gidro tany an-tsena)
 that PAST.saw lemur LOC PREP-market
 ‘Rabe says that he (Rabe) saw a lemur at the market.’
- b. Marary (satria nihinana voankazo manta) Rasoa
 sick because eat fruit unripe Rasoa
 (satria nihinana voankazo manta)
 because eat fruit unripe
 ‘Rasoa is sick because she (Rasoa) ate unripe fruit.’

We propose that the obligatoriness of CPEX only when the CP has an overt clause-final subject follows from the prosodic structure of the language and a principle barring certain kinds of prosodic recursion. To see this, we need to develop a basic picture of Malagasy declarative clause intonation.

We assume a prosodic hierarchy consisting of intonational phrases (ι), phonological phrases (φ), and phonological words (ω) (Nespor and Vogel 1986, Selkirk 1986). Following Selkirk 2009 and Bennett et al. to appear, prosodic structure is read off of the syntactic structure using the default Match principles in (36).

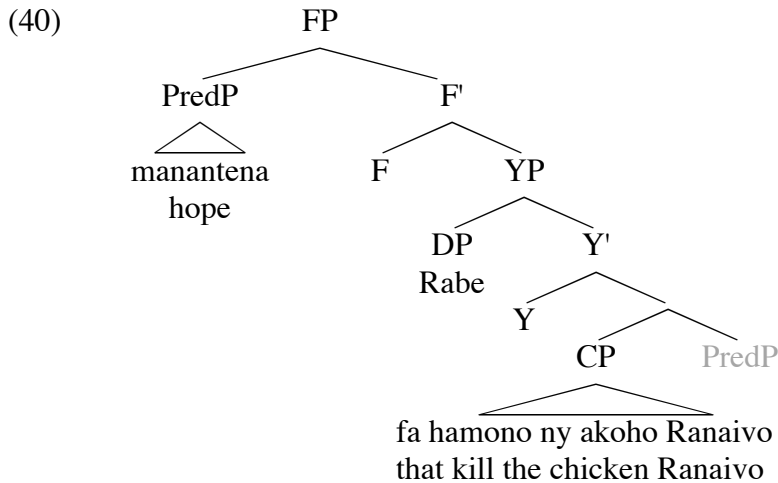
- (36) a. Match Word: Phonological words (ω) correspond to heads of phrases
 b. Match Phrase: Phonological phrases (φ) correspond to phrases
 c. Match Clause: Intonational phrases (ι) correspond to clauses

A VOS clause in Malagasy with the predicate fronting structure repeated in (37) has the prosodic structure in (38) (ignoring prosodic structure below the phonological phrase). This structure accords with descriptions of Malagasy intonation, which clearly identify the subject and predicate in a Malagasy clause as constituting separate phonological phrases (Dahl 1952, 1996, Rafitson 1980, Raoniarisoa 1990). Frascarelli 2010 indicates that the predicate can be identified by a rising tone on its last word, aligned with the stressed syllable (L*+H using the ToBI labelling system). There is a secondary rising tone on the subject.

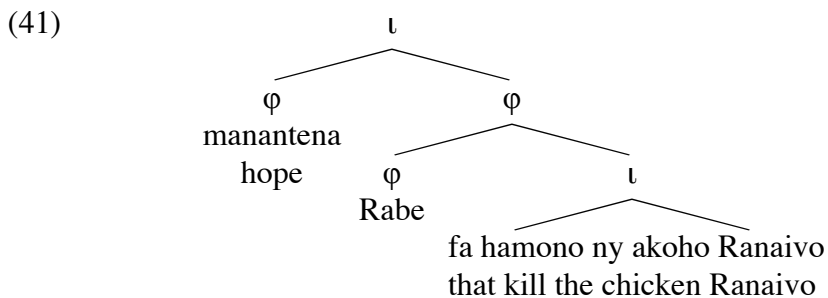


Adopting this picture, we can begin to explain the restriction on CPEX. Consider first a grammatical example of CPEX in (39). This sentence has the structure in (40).

- (39) Manantena Rabe **fa hamono ny akoho Ranaivo**
 hope Rabe that kill DET chicken Ranaivo
 ‘Rabe hopes that Ranaivo will kill the chicken.’



Given the Match principles, its prosodic structure should be as follows:

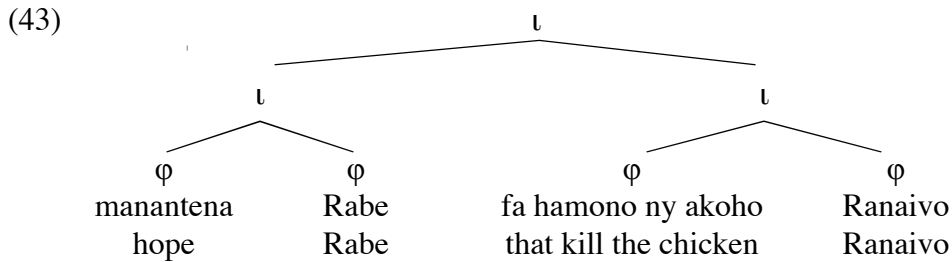


The structure in (41) is ill-formed however. As Bennett et al. to appear:60-62 discusses, certain kinds of recursion are not possible in prosodic structure. In particular, a higher-level phonological constituent may not contain a lower level one. That is, ι may contain ϕ but ϕ may not contain ι . We call this the Nested Layering Constraint, (42).³ Such a constraint differentiates prosodic structure from syntactic structure, which is not restricted in this way.

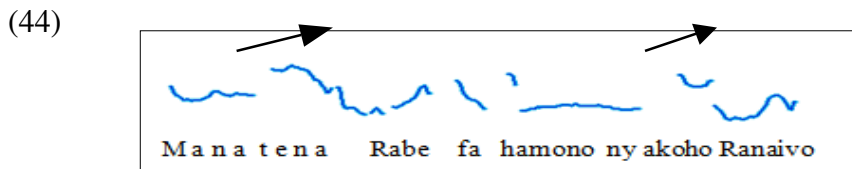
³ This constraint is reminiscent of the Strict Layer Hypothesis (Selkirk 1984, Nespor and Vogel 1986) but is less restrictive in allowing for recursion. Recent work strongly suggests that recursion of like prosodic constituents should be allowed (Ladd 1986, Wagner 2010, Elfner 2012, others).

- (42) **Nested Layering Constraint (NLC)**
 A prosodic constituent of type A may not contain a constituent B that is higher on the prosodic hierarchy

In order to avoid violating the NLC, the prosodic structure in (41) is adjusted (Bennett et al. to appear): the intonational phrase corresponding to the complement clause is adjoined to the matrix intonational phrase, yielding a well-formed representation, (43).

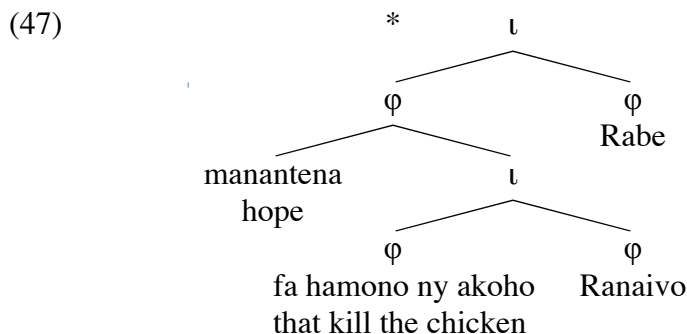
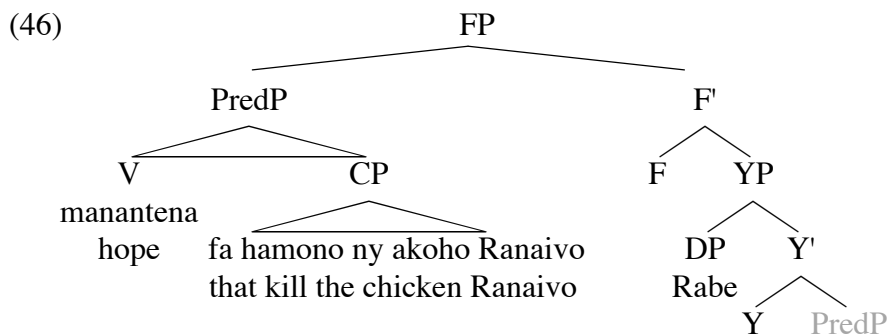


This analysis is confirmed by the pitch track for the sentence in (44). The rising tones on the phonological phrases corresponding to the two predicates are shown. One can also see the secondary rising tones on the subjects.



Now consider the ungrammatical example without extraposition of the CP, in (45). Given the syntactic structure in (46), the predicted prosodic structure is in (47). The embedded CP constitutes its own intonational phrase in accordance with the Match principles. The matrix clause in turn consists of an intonational phrase containing the predicate ‘hope that Ranaivo will kill the chicken’ and the subject ‘Rabe’, as shown.

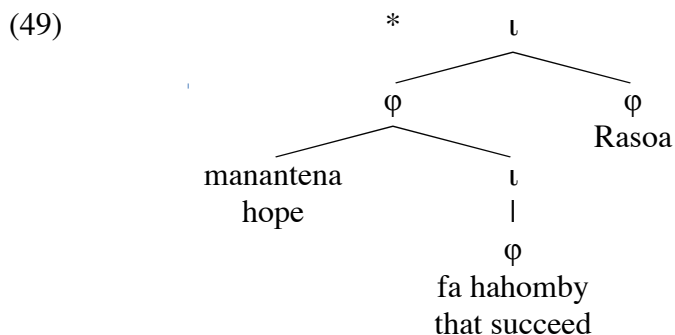
- (45) *Manantena **fa hamono ny akoho Ranaivo** Rabe
 hope that kill DET chicken Ranaivo Rabe
 (‘Rabe hopes that Ranaivo will kill the chicken.’)



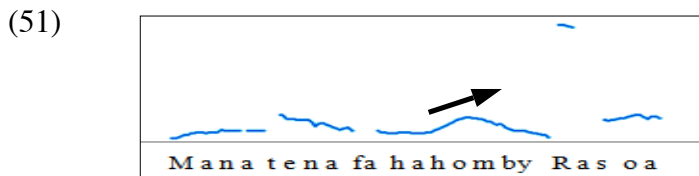
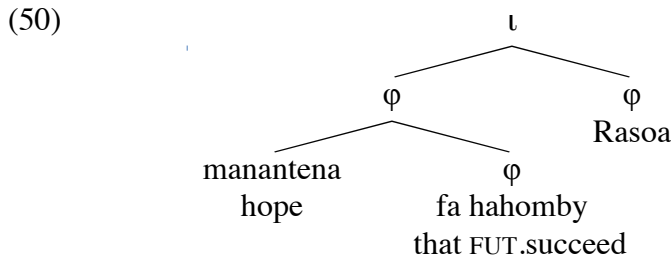
This structure also violates the NLC. It can be adjusted as discussed above to yield the representation in (43). Crucially, however, this would change the word order and would not correspond to (45), but (39). Either the CP is extraposed in the syntax or its position is phonologically adjusted in the prosodic structure. Either way, the word order that results is (39). (45) is never generated.

Finally, we consider an example of optional extraposition when the clause-final subject is absent, as in the topic-drop example in (48). The predicted prosodic structure using the Match principles is (49) (compare to (47)).

- (48) Manantena **fa** **hahomby** e_i Rasoai
 hope that succeed Rasoai
 ‘Rasoai hopes that she (Rasoai) will succeed.’



This structure violates the NLC. One way to rescue it is to extrapose the CP, but that would not yield the desired word order. An alternative suggests itself, appealing to binarity: an optimal prosodic constituent contains exactly two subconstituents (Inkelas and Zec 1990). The embedded intonational phrase in (49) does not contain two phonological phrases. This is a direct consequence of the complement clause missing a subject, which would constitute a second phonological phrase. We hypothesize that a more optimal representation eliminates ι , yielding the representation in (50), which is well-formed. It corresponds to the pitch track for this sentence in (51), where there is only one predicate-related high tone at the end of the embedded clause.



In summary, we believe that the obligatoriness of CPEX follows from general and language-specific prosodic requirements. Consequently, CPEX can be viewed as an instance of simple extraposition, its obligatoriness a result of these independent restrictions.⁴

5. Conclusion

Extraposition in Malagasy is a commonplace phenomenon in which predicate-internal elements appear clause-finally, after the subject. It is generally optional except that it is impossible for objects and obligatory for clauses. We proposed that this obligatoriness is due to a prosodic restriction against recursion in which an intonational phrase is embedded inside a phonological phrase. The illicit recursion can be resolved through extraposition, or by removing the embedded clause subject, which reduces the intonational phrase to a phonological phrase. Thus, the

⁴ The obligatoriness of CPEX strikes us as the reason why CPEX overrides the backgrounding function of extraposition. Without this override, there would be no way to express certain information structures. For example, it would not be possible to assert the propositional content of a CP. We save for later a formal implementation of this intuition.

obligatoriness of CP extraposition has an independent account and is not a sufficient reason to recognize two kinds of extraposition, at least not for Malagasy (contra Manetta 2012 for Hindi).

Extraposition has a semantic consequence of backgrounding the extraposed constituent but obligatorily extraposed CPs do not need to be interpreted as backgrounded.

Syntactically, we argued that extraposed constituents are not base-generated in the extraposed position but are A' moved there. They behave syntactically as though they are in the predicate-internal position as evidenced by a variety of reconstruction facts.

A number of issues require further investigation. We quickly mention three: First, why can objects not extrapose? Second, Malagasy has other kinds of extraposition: Extraposition from NP and Heavy XP Shift. Do they have the same analysis? Finally, why does obligatory CPEX not have the same backgrounding function as optional extraposition? We hope to address these questions and others in future work.

References

- Baltin, Mark. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13:1-38.
- Baltin, Mark. 2006. Extraposition. *The Blackwell companion to syntax*, eds. Martin Everaert and Henk van Riemsdijk, 237-271. Oxford: Blackwell.
- Bennett, Ryan, Emily Elfner, and James McCloskey. to appear. Pronouns and prosody in Irish. *Selected Proceedings of the XIV International Congress of Celtic Studies*.
- Bruening, Benjamin. 2010. Ditransitive asymmetries and a theory of idiom formation. *Linguistic Inquiry* 41:519-562.
- Büring, Daniel and Katarina Hartmann. 1997. Doing the right thing—Extraposition as a movement rule. *The Linguistic Review* 14:1-42.
- Dahl, Otto Christian. 1952. Étude de phonologie et de phonétique malgache. *Norsk Tidsskrift for Sprogvidenskap* 16:148-200.
- Dahl, Otto Christian. 1996. Predicate, subject and topic in Malagasy. *Oceanic Linguistics* 35:167-179.
- Elfner, Emily. 2012. Syntax-prosody interactions in Irish. Doctoral dissertation, University of Massachusetts at Amherst.
- Fitzpatrick, Justin. 2005. The whys and how comes of presupposition and NPI licensing in questions. *Proceedings of the 24th West Coast Conference on Formal Linguistics*, eds. John Alderete et al., 138-145. Somerville, Ma.: Cascadilla Proceedings Project.
- Frascarelli, Mara. 2010. Intonation, information structure and the derivation of inverse VO languages. *Austronesian and theoretical linguistics*, eds.

The Proceedings of AFLA 22

- Raphael Mercado, Eric Potsdam, and Lisa Travis, 81-102. Amsterdam: John Benjamins.
- Göbbel, Edward. 2007. Extraposition as PF movement. *Proceedings of the 34th Western Conference on Linguistics*, eds. Erin Bainbridge and Brian Agbayani, 132-145. Fresno, Ca.: California State University.
- Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.
- Kalin, Laura. 2009. News about the *no*-phrase: Specificational pseudoclefts in Malagasy. Undergraduate thesis, McGill University.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Ma.: MIT Press.
- Keenan, Edward L. 1976. Remarkable subjects in Malagasy. *Subject and topic*, ed. Charles Li, 247-301. New York: Academic Press.
- Keenan, Edward L. 1995. Predicate-argument structure in Malagasy. *Grammatical relations: Theoretical approaches to empirical questions*, eds. Clifford S. Burgess et al., 171-216. Stanford: CSLI.
- Ladd, D. Robert. 1986. Intonational phrasing: the case for recursive prosodic structure. *Phonology* 3:311-340.
- Law, Paul. 2007. The syntactic structure of the cleft construction in Malagasy. *Natural Language and Linguistic Theory* 25:765-823.
- Manetta, Emily. 2012. Reconsidering rightward scrambling: Postverbal constituents in Hindi-Urdu. *Linguistic Inquiry* 43:43-74.
- Massam, Diane. 2000. VSO and VOS: Aspects of Niuean word order. *The syntax of verb initial languages*, eds. Andrew Carnie and Eithne Guilfoyle, 97-116. Oxford: Oxford University Press.
- Massam, Diane, and Caroline Smallwood. 1997. Essential features of predication in English and Niuean. *Proceedings of the 27th North East Linguistic Society*, ed. Kiyomi Kusumoto, 263-272. Amherst, Ma.: GLSA Publications.
- Nespor, Marina, and Vogel, Irene. 1986. *Prosodic phonology*. Dordrecht: Foris.
- O'Grady, William. 1998. The syntax of idioms. *Natural Language and Linguistic Theory* 16:279-312.
- Paul, Ileana. 2000. Malagasy clause structure. Doctoral dissertation, McGill University.
- Paul, Ileana. 2005. Or, *wh*-, and not: Free choice and polarity in Malagasy. *UCLA Working Papers in Linguistics 12: Proceedings of AFLA XII*, 359-367.
- Paul, Ileana, and Lucie Rabaovololona. 1998. Raising to object in Malagasy. *The structure of Malagasy, vol. II*, ed. Ileana Paul, 50-64. Los Angeles, Ca.: UCLA Department of Linguistics.
- Pearson, Matthew. 2001. The clause structure of Malagasy: A minimalist approach. Doctoral dissertation, UCLA.

- Potsdam, Eric. 2007. Malagasy sluicing and its consequences for the identity requirement on ellipsis. *Natural Language and Linguistic Theory* 25:577-613.
- Potsdam, Eric, and Maria Polinsky. 2007. Missing complement clause subjects in Malagasy. *Oceanic Linguistics* 46:277-303.
- Rackowski, Andrea, and Lisa Travis. 2000. V-initial languages: X or XP movement and adverbial placement. *The syntax of verb initial languages*, eds. Andrew Carnie and Eithne Guilfoyle, 117-141. Oxford: Oxford University Press.
- Rafitoson, Elisa. 1980. Accent et intonation de la phrase " objective" en malgache-merina. Doctoral dissertation, Université de Provence.
- Rajaonarimanana, Narivelo. 1995. *Grammaire moderne de la langue malgache*. Paris: L'Asiathèque.
- Raoniarisoa, Noro. 1990. Accent and intonation in a Malagasy dialect. Doctoral dissertation, University College of North Wales.
- Rochemont, Michael, and Peter Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT.
- Selkirk, Elizabeth. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge, Ma.: MIT Press.
- Selkirk, Elizabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3:371-405.
- Sheehan, Michelle. 2010. Extraposition and antisymmetry. *Linguistic variation yearbook* 10, ed. Jeroen von Craenenbroeck, 201-251. Amsterdam: John Benjamins.
- de Vries, Mark. 2009. Specifying coordination: An investigation into the syntax of dislocation, extraposition, and parenthesis. Manuscript, University of Groningen.
- Wagner, Michael. 2010. Prosodic recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory* 28:183-237.
- Wilder, Chris. 1996. Rightward movement as leftward deletion. *On extraction and extraposition in German*, eds. Uli Lutz and Jürgen Pafel, 273-309. Amsterdam: John Benjamins.
- Winterton, Matthew. 2011. Malagasy-English phrasebook and dictionary.
- Zribi-Hertz, Anne, and Liliane Mbolatianavalona. 1999. Towards a modular theory of linguistic deficiency: Evidence from Malagasy pronouns. *Natural Language and Linguistic Theory* 17:161-218.

ON THE MARSHALLESE SEMI-GLIDES AND UNDERSPECIFIED
VOWELS*

Catherine Provencher
Université du Québec à Montréal
catherineprovencher@hotmail.com

This paper attempts to refine the set of Marshallese phonemes by reducing the number of underspecified vowels. We will propose that even though Marshallese has a vertical vowel system, some surface forms may actually be the result of different phonological processes. The theory of elements will be used to uncover the phonological primes of this particular language and describe their interaction as well as the constraints in which they operate. An ulterior motive is to provide evidence that phonology as a part of language share basic properties with other components of the grammar such as asymmetrical relationships between primitives.

1. Introduction

Marshallese (henceforth MRS) is a language of the Micronesian family whose phonology has been extensively studied over the last few decades. It has been suggested that MRS possesses a vertical vowel system in which the quality of the vowel is dependent on the consonants' secondary places of articulation (Bender 1968). This proposal entails that there is a set of underspecified vowels, either three (Bender 1968, Choi 1992) or four (Abo et al. 1976, Hale 2000; 2007), and these are specified for height (Bender 1968, Choi 1992, Abo et al. 1976) or height and ATR (Hale 2000; 2007). It follows that backness and rounding features are provided by the neighbouring consonants. However initial and final vowels, as well as long medial vowels, do not always conform to the aforementioned rule: it is not the case that the vowel reflects the features of the consonants and thus Bender (1968) proposes that there are also semi-glides, namely j-w-h, presumably inside the nucleus that are responsible for the final phonetic output. Regardless there is a lack of consensus regarding the actual set of phonemic surface forms that constitute the MRS vowel inventory. We will propose that there are only two underspecified vowels and that the presence of those semi-glides is not required in all

*This work is supported in part by funding from the Social Sciences and Humanities Research Council of Canada to the Major Collaborative Research on Interface Asymmetries, grant number 214-2003-1003, and by funding from the FQRSC on Dynamic Interfaces, grant number 137253. www.interfaceasymmetry.uqam.ca. www.biolinguistics.uqam

circumstances if at all. The purpose of this paper is twofolds: to clarify the phonemic and phonological properties of MRS and, most importantly, to show that phonology as a part of the language faculty must possess the same properties and obey the same principles as morphology or syntax. As Di Sciullo (2005: 21) puts it: “The components of the grammar share basic properties and differ with respect to their primitives, operations, and conditions.” We will first focus on the identification of phonological primitives using the Element Theory framework (Kay, Lowenstamm & Vergnaud 1985, Harris 1994, Backley 2011 among many others) within which each phonetic property such as place, manner, voicing, etc. corresponds to a unipolar element. We will particularly focus on the elements used in the characterization of vowels. The second part will be concerned with the relations holding between these primitives. We will assume that these relations are asymmetrical i.e. that elements are combined in a governor/dependent relation (Clements 1985, van der Hulst 1988; 1989, Harris 1994). Finally we will attempt to explain the conditions in which these relations take place in MRS.

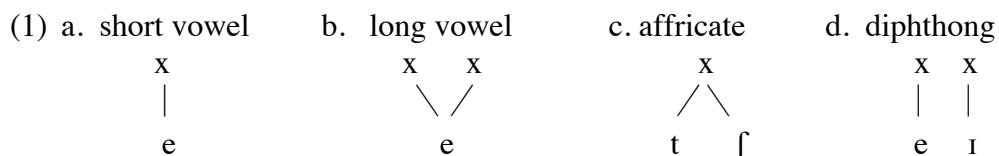
2. Phonological Primes.

Phonemes interact with each other such that the final phonetic output of specific words may differ depending on their environment. In order to describe such processes we must first know the exact components of phonemes responsible for such processes. Since *Sound Pattern of English* (Chomsky & Halle 1968) phonology is viewed as a set of rules applied to phonological segments which are themselves conceived as bundles of features. Those features describe manner and place of articulation as well as voicing and are represented in a binary fashion with either + or - whichever feature. Segments and their respective set of features are arranged successively on a string hence the term linear phonology. There are however other frameworks that are considered non-linear such as element theory that is part of the broader auto-segmental framework.

2.1 Auto-segmental Phonology.

Non-linear theories of phonology are not concerned with linear representations but rather views segments as skeletal slots associated with elements arrayed on independent tiers (Goldsmith 1979). Each phoneme is represented by a position x that is associated with phonological quantity i.e. the comparative duration between different segments. Each x is positioned on a skeletal tier or skeleton and, with the help of association lines, is linked to the segment that represents the phonetic output. A phoneme may occupy two skeletal slots doubling the duration of pronunciation. Long vowels and geminates are good examples of this. It is also possible to have two phonemes associated with one timing slot as is the case for affricates which are quantitatively more complex. Finally diphthongs, although

considered as being one nucleus, will nevertheless be characterized by two different timing slots each associated with one part of the nucleus:



Since phonology is the interaction between phonetic segments, phonological processing is basically association or dissociation between segments and features. On the one hand a feature can spread to an adjacent slot changing the nature of the segment: this is the assimilation process. On the other hand a segment can lose its association with a given feature: this is dissimilation. The features mentioned here are not exactly the same type of features found in linear phonology but rather elements that are « fully specified feature matrices » (Kaye, Lowenstamm & Vergnaud 1985 : 311).

2.2 Element Theory

In this framework each phoneme may contain one or several elements embodying different characteristics. These elements are considered the primitives of phonology. Many lists of elements with different interpretations have been proposed. However our prime focus being vowels we will concentrate on the particular elements that comprise them. Those are [I], [U], [A], and [@]. According to Anderson & Jones (1974) [I] represents ‘frontness’, [A] represents ‘lowness’ and [U] represents roundness. We will however use the more familiar terms height and backness for [A] and [I] respectively. [@] is called the neutral element or the cold vowel and according to Harris (1994: 109) it stands for the « neutralization of peripheral vowel qualities under a centralized reflex ». In other words it gives a vowel a schwa-like quality that some might interpret as the tense-lax duality. We will not discuss much the cold vowel in this analysis since other means of deriving lax vowels seem more economical to us.

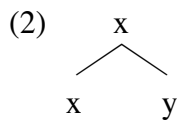
We are thus left with three different elements that may also be part of consonants. According to Harris (1994) these can also represent a place of articulation such that a consonant containing [I] would be palatal, a consonant with [U] a labial, a consonant with [A] would be uvular, and a consonant with [@] is a velar. Van der Hulst (1988; 1989) as well as Backley and Nasukawa (2009) propose that [U] may actually represent both labial and velar. With the consequence that we can completely eliminate the cold vowel from our inventory.

Furthermore it is also proposed that [A] may exemplify a pharyngeal quality. In terms of representation it seems that elements representing vocalic

features may possess two tiers each depicting different articulatory gestures (van der Hulst 1988, 1989). In other words elements may be complex entities.

3. Asymmetry in Phonology

An operation of the language faculty is to associate two items to form a new one. In syntactic theory (especially within the minimalist program) this operation is called merge (see Chomsky 1993, 1995, 2000 for more details). These items are arranged in an asymmetrical fashion that is, one as precedence or dominance over the other. The same property is found in morphological items as well (see Di Sciullo 2005). In any case the established representation of such asymmetries is in the form of a tree where a head precedes a dependent and also projects so that it may govern its dependent as well:



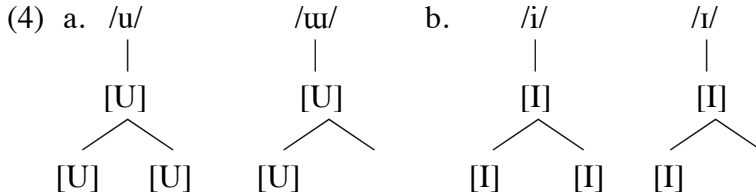
3.1. Asymmetrical Properties of Phonological Primes

If we assume that each element has this basic structure then each position must represent a different phonetic property of a given element. According to van der Hulst (1988, 1989) the governor position represents the constriction place during the production of a vowel. The dependent on the other hand is associated with a secondary manner:

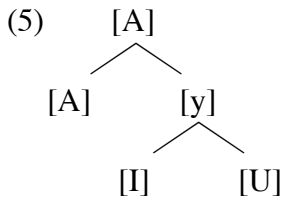
- (3) a. [A] governor: pharyngeal constriction
dependent: openness of the jaw
- b. [I] governor: palatal constriction
dependent: advanced tongue root
- c. [U] governor: velar constriction
dependent: rounding of the lips

A consequence of associating two different phonetic interpretations for the same phonological element is the reduction of primitives which is in accordance with the principle of economy. The element [A] as a governor yields a low vowel and an open vowel when it is a dependent. The element [U] produces a back vowel as a governor and a rounded one as a dependent. Finally the element [I] yields a front vowel as a governor and an ATR vowel as a dependent. One must assume that a governor may select the same element as a dependent. This is what van der Hulst

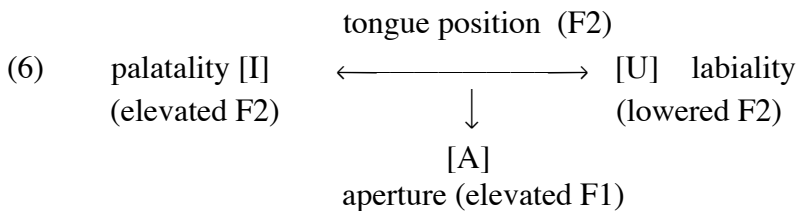
(1988, 1989) calls the Universal redundancy rule. Evidently a governor that does not select any dependent will yield a different phonetic output:



Since there are three available primitives for the composition of vowels and since one of the properties of language is its binarity (i.e. there are no syntactic tree with three branches), Ewen & van der Hulst (1988) suggest a two-way split for the vocalic elements that captures the precepts of asymmetry:



Based on their work on harmony systems the element [A] is opposed to [I] and [U] in that the election of these latter as governors might be neutralized in some languages. Therefore the laryngeal property is opposed to what they call the ‘tongue body constriction’ represented by /y/. This is consistent with the acoustic structure of elements (from Shane 2005: 338):

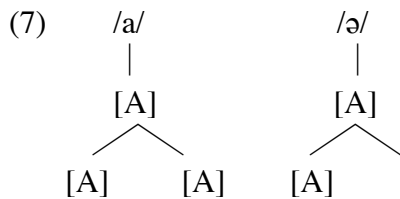


The aperture (opening of the jaw) represented by [A] contrast with the tongue position indicated by the second acoustic formant (F2). An elevated F2 indicates a tongue displaced to the front whereas a low F2 represents a back displacement of the tongue. According to Wood (1979) the tongue is drawn rearwards during the production of [u]-like classes of vowels. The velum is also tightly closed. This is in accordance with the assumption that [U] characterizes the velar constriction of a back vowel. Moreover the rounding of a vowel is expressed by a third acoustic formant (F3) providing evidence for the double interpretation of

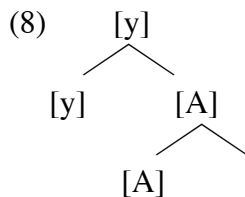
the [U] element. Since back unrounded vowels are rare amongst the languages of the world the association of both properties in one single element is straightforward (van der Hulst 1988). However we must account for the fact that this type of restrictions is not universal i.e back unrounded vowels do exist. The question is thus how to explain such variation.

3.2 The Completeness of Elements

We have stated that elements have different properties depending on their structural position. However there are also constraints that block the possibility of separating elements. This results in the absence of certain vocalic phonemes in a given language. For example a language with a complete [U] cannot possess back unrounded vowels. A language with a complete [I] does not possess front -ATR vowels, or as van der Hulst calls them, central vowels. An incomplete [A] represents a low vowel that is not completely opened i.e /ə/. However when [A] is complete then it can only represent a low vowel that is -ATR and -round namely /a/. Therefore the structures of [A] are as follows:



A complete [A] in a given language would result in the absence of a schwa. To recapitulate, the completeness or incompleteness of an element is language-specific. This variation can be illustrated by the use of a double skeletal tier each representing a feature. The internal structure of elements should also express this duality depending on the number of features expressed. One must remember however that [A] is treated apart from the other two elements. According to van der Hulst (1989) when a segment comprises [A] plus another element it is automatically considered [+closed]. Therefore it is impossible for [A] to be the dependent unless it is incomplete:



Vowels with [A] as a dependent can only refer to mid or higher vowels that are closed. The value of [y] can either be [I] or [U] whether complete or not: a segment can retain the [+ATR] or the [+round] features without these being associated with [+front] or [+back] respectively.

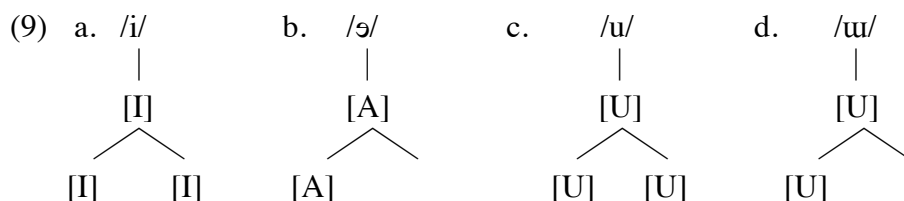
4. Analysis of MRS Primes

The first step in analyzing MRS data is to outline the properties of its primes. Let us start by delineating the nine phonemes that results from our assumption that there are no lax vowels in MRS unless they are the result of some sort of phonological processes (see Bender 1968 for more details):

TABLE 5: TENTATIVE MRS VOWEL INVENTORY

	Front	Back	
		Unrounded	Rounded
High	i	ɯ	u
Mid	e	ɘ	o
Low	æ	ɑ	ɒ

We can observe in Table 5 that MRS has no central vowels barring the schwa. This indicates that [I] is complete and therefore cannot be used as a dependent resulting in a +ATR vowel. This is consistent with the fact that ATR is not a relevant feature in MRS phonemics. Consequently the use of [I] in our analysis will always indicate a front vowel that is always +ATR. The presence of the schwa as a central vowel suggest that [A] on the other hand is not complete and that in fact it is never complete in MRS due to the absence of /a/ as a possible phoneme. Finally [U] may also be incomplete since there are back unrounded vowels in MRS. In sum, these are the possible MRS primes (with the corresponding phonetic outputs):



Only these primes can be combined together to produce novel phonemes. However there is also a restriction regarding which prime can be combined with one another. Notice that in MRS there are no front rounded vowels. This

denotes the impossibility to combine [I] with [U]. Therefore (9 a-c-d) can only merge with (9b).

4.1 MRS Phonological System

Since Bender (1968) MRS vowel system has been treated as a vertical system meaning that there are only a few underlying vowels that surfaces as phonetic outputs only after being influenced or colored by the consonantal environment. In the case of MRS it is the consonant's secondary places of articulation that are responsible for the quality of the vocalic phonemes. Consonants can either be palatalized, velarized or rounded. According to the theory of elements this would mean that palatalization is provided by [I], velarization by incomplete [U], and rounding by complete [U]. In a #CVC# environment where both consonants have the same secondary place of articulation it is always the case that the vowels will be front between palatalized consonants, back unrounded between velarized consonants and back rounded between rounded consonants (from MOD)¹:

- | | | | |
|------|---|--|--|
| (10) | C^jVC^j | C^yVC^y | C^wVC^w |
| | m ^j it ^j 'dead, numb' | k ^y ur ^y 'dig' | k ^w uk ^w 'gathered together' |
| | t ^j el ^j 'curved' | p ^y ək ^y 'bring' | l ^w oŋ ^w 'a fly' |
| | p ^j ælj 'to die' | p ^y at ^y 'slow' | l ^w oŋ ^w 'ant' |

This is not the case however when the environment is asymmetrical; the quality of the vowel is not predictable except when one of the consonant is rounded, in that case the vowel is always round. Therefore the problematic data always involves a palatalized and a velarized consonant (from MOD):

- | | | |
|------|---|--|
| (11) | C^jVC^y | C^yVC^j |
| | t ^j ip ^y : 'rise, of bread' | kit ^j : 'bacteria' |
| | p ^j et ^y : 'foul' | kəl ^j : 'fertile soil' |
| | n ^j əl ^y : 'make weapons' | l ^y am ^j : 'bay' |
| | l ^j al ^y : 'world' | |

The first thing to notice is that /e/ only occurs when the first consonant is palatalized. Moreover /æ/ never appears in asymmetrical environments. We have said that in MRS [I] and [U] never merge since there are no front rounded vowel. Suppose then that this is also the case when these elements are competing to spread

¹ All examples are taken from a survey of all #CVC# words in the Marshallese Online Dictionary (MOD). Loanwords, most of which are from English, were excluded. Longer words were excluded as well on the basis that there are stress patterns that may influence the vocalic outputs. Although there are words containing lax vowels in the #CVC# format, we have purposefully ignored them for the time being.

to the same nucleus. We may assume that [I] is blocked by [U], as is the case in the presence of a rounded consonant, and thus the vowel that surfaces is back unrounded /ɑ/ instead of front unrounded /æ/. The case of /e/ is a little more problematic, while /æ/ and /ɑ/ have [A] as their head, /e/ uses this element as a dependent. So it seems that only the preceding consonant may spread to the following nucleus, provided that the latter has an empty head i.e has no governor.

This raises the fundamental question on the nature of the nucleus and what occupies this position before any consonantal influence. In a vertical vowel system, these are the underlying vowels specified only for height. We will propose that there are only two underlying vowels namely [A] and [...]. In (9) we stated that there were four phonological primes in MRS and if we combine them together, we get the complete set of MRS nine surface vowels. However only one prime could actually be merged with the others: incomplete [A] in (9b). We will therefore assume this is the first underlying vowel. The other primes are basically the secondary places of articulation that have spread to an empty nucleus: /i/ is the result of [I] spreading, /u/ is the result of complete [U] spreading, and /ʊ/ is the result of incomplete [U] spreading. The final problem to address is why sometimes [A] is the governor as in /æ/, /ɑ/, and /ɒ/ and sometimes is the dependent as in /e/ and /o/. The answer might be found from a diachronic perspective.

4.2 Proto-Micronesian Phonology

The ancestor of MRS had a different phoneme inventory. The most striking point is the lack of secondary places of articulation except for p^w and m^w which are lost in MRS². Moreover PMC had only five vowels: *i, *e, *a, *o, *u. We will propose that MRS consonant's secondary places of articulation come from PMC vowels that lost the element that was in their governor position. For *i and *e [I] spread to the preceding consonant, for *u and *o, [U] spread to the preceding consonant. What was left behind was either an empty nucleus, or incomplete [A] as a dependent. This process gives us our two underlying vowels. However PMC *a had a distinctive behavior; no element could spread to the preceding consonant. Therefore it remained a complete [A] that resulted in the low vowels of MRS namely /æ/, /ɑ/, and /ɒ/ (from Bender et al. 2003):

- (12) a. *kani > kan^j 'eat s.t.
*lama > l^yam^j 'bay'
*kara > kar^y 'scrape'

² Phonemes or words that are preceded by an asterisk are reconstructed forms.

- b. *tañi > tʰænʲ ‘from (s.t.), away from’
*mata > mʲætʲ ‘eye, face’
*papa > pʲæpʲ ‘coconut frond, midrib of frond’
- c. *laño > lʷɒŋʷ ‘a fly’
*lako > lʷɒkʷ ‘toward a third party,’
*maño > mʲɒŋʷ ‘pate of head, fontanelle’

As we can observe in (12) most PMC words have lost their final vowels in MRS. According to Hale (2007) this is the result of a penultimate stress: CVC-V > CVC. In (12a) PMC *a > ʌ in the presence of either a velar or a velarized consonant; in (12b) PMC *a > æ in between two palatalized consonants and in (12c) PMC *a > ɒ in the presence of a rounded consonant. In most circumstances lost final PMC *a velarizes a MRS consonant. However the PMC consonant inventory was significantly reduced over time: fricatives and palatals, as well as some dentals are lost in MRS. The result of this loss was either a palatalized version of the old consonant or a complete lost. The palatalized versions are impervious to the spreading from a neighboring vowel. Therefore even if a PMC word ended with *a the previous consonant may not be velarized in MRS. For example PMC */camʷa/ > MRS /rʲamʷ/: the palatal *c became a palatalized /r/. In (12b) PMC *papa should have become MRS /pʷapʷ/ with velarized consonants. But the contrast in PMC between the labials /p/ and /m/ on the one hand and the labio-velars /pʷ/ and /mʷ/ on the other survived in MRS: the labials became palatalized and the labio-velars became velarized yielding the form /pʲæpʲ/. Thus two different diachronic consonantal changes may have had an influence over MRS surface vowels. We will now propose that the set of semi-glides Bender (1968) proposes to account for the lack of consonantal influence in some specific environments are in fact remnants of PMC.

4.3 The semi-glides

We have said that MRS vowels are always colored by the neighboring consonants. But when it comes to word initial, word final and long middle vowels, this is not always true. We may find a round or a front vowel next to a velarized consonant or a back vowel next to a palatalized consonant. To account for this Bender (1968) proposes a set of three semi-glides namely j-w-h. However Choi (1992; 1995) finds no phonetic evidence for the presence of these semi-glides. Moreover in terms of auto-segmental theories those semi-glides cannot occupy their own slot. We will propose instead that the unexpected quality of the vowel may result from diachronic change. For example if a PMC word started with a front vowel, the [I] contained in that vowel had no place to spread since there were no preceding consonant. This [I] would remain in the governor position of that nucleus and

would therefore block any type of spreading yielding the unpredicted front vowel. Here are a few examples (from Bender et al. 2003):

- (13) a. *ate > atⁱ ‘liver’
 *ale > alⁱ ‘song’
- b. *ika > ek ‘fish’
 *ira > er^y ‘they, ABS’
- c. *una > unⁱ ‘fish scale’
 *ura > or^y ‘lobster’

In (13a) a back vowel is followed by a palatalized consonant, in (13b) a front vowel is followed by a velarized consonant and in (13c) a round vowel is followed by either a palatalized or a velarized consonant. We cannot assume the same process regarding final vowels since these had a previous consonant to which an element could spread. However it is the case that sometimes the consonant changed before spreading could happen and thus the final vowel remained colored as it was in PMC (from Bender et al. 2003):

- (14) *cuu > rⁱu ‘gather to dance’
 *fitu > itⁱu ‘star’
 *itau > tⁱit^o ‘*Callophyllum inophyllum*’
 *katuu > katⁱu ‘mast’
 *kawakawa > koko ‘dolphin’

The PMC words in (14) all contained a final rounded vowel (or a /w/ in the case of *kawakawa) and that final vowel is still present in MRS. Systematic change happened to the preceding consonants such that *c > rⁱ and *t > tⁱ in MRS. These became impervious to the spreading of the final vowel governor element. Therefore this element remained in the final nucleus. The final example in (14) may be explained by the fact that /a/ and /w/ merged to produce /o/. This is of course a very small sample but it seems that there are not many examples whose final vowel does not correspond to the appropriate coloration. Other examples may be found in English loan words (from MOD):

- | | | | | |
|------|----------------|---|-----------------|------------------------------------|
| (15) | <i>antenna</i> | > an ^t y ^e n ⁱ a | <i>corner</i> | > kon ⁱ a |
| | <i>purser</i> | > p ^y at ⁱ a | <i>to marry</i> | > m ^y ar ^y e |
| | <i>bottle</i> | > p ^y at ^y o | <i>razor</i> | > r ^y et ⁱ a |
| | <i>care</i> | > kea | <i>tally</i> | > t ^y al ^y e |
| | <i>Gehenna</i> | > ken ⁱ a | <i>hose</i> | > oot ⁱ |
| | <i>cue</i> | > kiu | <i>oak</i> | > uuk |

It seems like MRS will adopt the nearest sound possible in terms of pronunciation regardless of the intereaction between vowels and consonants. Moreover certain English consonants that do not exist in MRS will be systematically translated to the same MRS counterpart despite the environment. For example some fricatives and affricates have turned into palatalized coronal /tʃ/:

- | | | |
|------|-------------------------|----------------------------------|
| (16) | <i>siren</i> > tʃairʲiŋ | <i>examination</i> > tʃenʲætʲinʲ |
| | <i>size</i> > tʃaitʲ | <i>zipper</i> > tʃipʲɔrʲ |
| | <i>judge</i> > tʃætʲ | <i>change</i> > tʃænʲitʲ |

This is the same phenomena we observed from PMC to MRS. This suggests that variation processes operate in the same fashion whether synchronically or diachronically. Long medial vowels may also show an unexpected coloration. According to Bender (1968) there are not many words in MRS that contain long vowels and most of them are loanwords from English. It is quite straightforward then that many loanwords kept the same phonetic outputs or the nearby equivalent of all their phonemes regardless of the environment (from MOD) :

- | | | |
|------|-------------------------------|-----------------------|
| (17) | <i>pontoon</i> > pʲaanʲtʲuunʲ | <i>shoes</i> > tʃuutʲ |
| | <i>boat</i> > pʲootʲ | <i>June</i> > tʃuunʲ |
| | <i>box/fork</i> > pʲɔŋk | <i>gold</i> > koolʲ |
| | <i>volcano</i> > pʲoolʲkenʲo | <i>goat</i> > kootʲ |

As can be seen in (17) a round vowel from an English word will remain round despite the fact that the surrounding consonants may not be rounded. We may conclude then that there are no semi-glides coloring the vowels appropriately. Even Bender (1968) mentioned that they might be part of a deeper level of representation which is actually consistent with our analysis providing that an element may be floating, awaiting to reattach to an empty slot. In the case of PMC to MRS it is true that elements that are prohibited to attach to the onset slot reattached to their previous slot i.e. the nucleus. This process and the conversion of consonants into the correspondent MRS form (either from PMC or English) are responsible for the lack of agreement between vowel quality and consonant's secondary places of articulation.

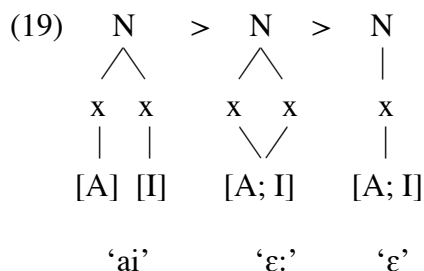
4.4 Lax vowels

At the beginning of section 4 we have deliberately excluded the set of lax vowels from our tentative inventory of MRS canonical vowels. Our primary motivation was the need to reduce the vocalic heights to three as per Bender (1968). His own exclusion of mid-high vowels on the ground that they were the result of phonological processes not directly part of a vertical vowel system was also quite convincing. Our assumption was that mid-low vowels were subject to the same

restrictions as mid-high vowels. In the dictionary of Abo et al. (1976) there are four vowel heights with the low and the low-mid vowels being contrastive i.e. words containing either one of these vowels may have a different semantic interpretation. In our review of all #CVC# pairs there are rare occurrences of such minimal pairs (from MOD):

- (18) a. p^jɛt^y : ‘foul’
 p^jɛt^y : ‘pillow’
- b. t^jɛl^j : ‘curved’
 t^jɛl^j : ‘grown over’
- c. k^wɔr^y : ‘crunch, rattle’
 k^wɔr^y : ‘frightened’

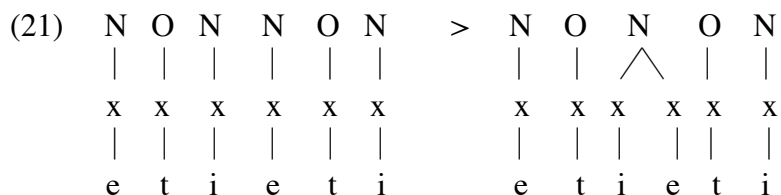
None of them is a loanword and they are exactly in the same environment. Although we have not found their PMC ancestor we will nevertheless propose that these different outputs are the result of a diachronic phonological process of diphthong reduction. Before we get to this we need to explain a few things about lax vowels. In terms of element theory, what differs between a tense and a lax vowel is that the latter has no governor. For example /e/ and /æ/ have the following elemental composition respectively: [I,A] and [A,I] with the underlined element being the governor. However /ɛ/ is the combination of [I] and [A] which stand in a symmetrical relationship. If PMC diphthongs were, at some point in time, shortened into a single phoneme, then the result of this process could be illustrated as this:



This process is attested in English although the vowel did not shorten (see Harris, 1994, for more details). Up to this point we have only found two such examples in PMC and quite interestingly they both involve reduplication (from Bender et al. 2003):

- (20) a. *etieti > ɛt^jɛt^j ‘to peel’ b. itaita > ɛt^jɛt^j ‘build up’

Reduplication is a common process in Micronesian languages and is still very productive in MRS. For the examples in (20) we may assume that the apparent diphthong was at first two different phonemes occupying each a different slot:



Then they were reanalyzed as being part of the same nucleus as a diphthong that was finally reduced to a short vowel. None of the element constituting this new vowel took precedence and thus it has no governor, yielding a short lax vowel. The number of reduplicated forms involving adjacent vowels is pretty scarce, at least in PMC. In any case, this hypothesis needs to be supported by more data yet it brings to light some more phonological processes that may be involved in diachronic change and we might find evidence for this process in loanwords as well.

5. Conclusion

The goal of this paper was to refine the set of MRS phonemes by reducing the number of underspecified vowels. We have provided some evidence that MRS may have only two underspecified vowels namely [A] and [...]. The combination of both with all three secondary places of articulation produced the set of nine basic MRS vowels. The lax vowels as well as the vowels that are not predictable from their consonantal environments were explained by consonant change and floating elements that reattached to their previous nucleus. Both these processes may be applied to the synchronic as well as to the diachronic plane providing evidence that the operations inherent to language variation apply to all sorts of change. Moreover our second objective was to provide evidence that phonology as a part of language share basic properties with other components of the grammar such as syntax and morphology. Our main concern was to define the primitives or to put it differently, the phonological objects, and describe their properties as well as the operations governing them. Element theory is perhaps one of the best approach for such endeavors. One important feature of this theory is the asymmetrical relations that hold between primes. In this respect phonology is no different than other components of the grammar.

References

- Abo, et al. 1976. *Marshallese-English Dictionnary*. Honolulu, University Press of Hawaii.
- Abo, Takaji. 2009. Marshallese-English Online Dictionary. University Press of Hawaii. Online <http://www.trussel2.com/mod/>
- Anderson, John & Charles Jones. 1974. Three theses concerning phonological representations. *Journal of Linguistics*, Vol. 10, No. 1: 1-26.
- Backley, Phillip. 2011. *An introduction to element theory*. Edinburgh: Edinburg University press.
- Backley, Phillip & Kuniya Nasukawa. 2009. Representing labials and velars: a single dark element. *Phonological Studies* 12: 3-10.
- Bender, Byron W. 1968. Marshallese Phonology. *Oceanic Linguistics*, vol.7: 6-35.
- Bender, Byron W. 1969. *Spoken Marshallese*. Honolulu: University of Hawaii Press.
- Bender, Byron W. et al. 2003. Proto-Micronesian reconstructions. *Oceanic Linguistics*. Vol. 42, no.1 : 1-110.
- Clements, G.N. 1985. The geometry of phonological features. *Phonology Yearbook* Vol. 2. 225-252.
- Choi, John D. 1992. Phonetic underspecification and target interpolation: An acoustic study of Marshallese vowel allophony. *UCLA Working Papers in Linguistics*, vol. 82.
- Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory Trad. de: *English*. Coll. MIT occasional papers in linguistics. Cambridge, MA: Distributed by MIT Working Papers in Linguistics.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass: The MIT Press
- Chomsky, Noam. 2000. Minimalist Inquiries: The framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, David Michaels and Juan Uriagereka Roger Martin (eds.) Cambridge: MIT Press. 89-155.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Di Sciullo, Anne-Marie. 2005. *Asymmetry in Morphology*. Cambridge, MA: MIT Press.
- Ewen, Colin & Harry van der Hulst. 1988. High, low, and back or [I], [A], and [U] ? In P. Coopmans & A. Hulk (eds.) *Linguistics in the Netherlands*. Dordrecht : Foris : 1-59.
- Goldsmith, John A. 1979. Autosegmental Phonology. PhD dissertation, MIT. New York : Garland.
- Hale, Mark. 2000. Marshallese phonology, the phonetics-phonology interface and historical linguistics.

The Proceedings of AFLA 22

- Hale, Mark. 2007. *Historical Linguistics: Theory and Method*. Oxford: Blackwell.
- Harris, John. 1994. *English Sound Structure*. Oxford : Blackwell.
- Hulst, H.G. van der. 1988. The geometry of vocalic features. In H. van der Hulst & N. Smith (eds). *Features, Segmental Structure and Harmony Processes*. Dordrecht : Foris. 77-126.
- Hulst, H.G. van der. 1989. Atoms of segmental structure: Components, gestures and dependency. *Phonology* 6/2, 253-284.
- Kaye, J., Lowenstamm & J.R. Vergnaud. 1985. The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook* 2. 305-328.
- Sanford A. Shane. 2005. The aperture particle /a/: Its role and functions. In Carr, Philip, Jacques Durand and Colin J. Ewen (eds). *Headhood, Elements, specification and contrastivity phonological papers in honour of John Anderson*. Amsterdam : John Benjamins. 313-338.
- Wood Sydney. 1979. A radiographic analysis of constriction locations for vowels. *Journal of Phonetics*. Vol.7 :25-43
- Zewen, F.X.N. 1977. *The Marshallese Language: A study of its phonology, morphology, and syntax*. PhD dissertation. University of Hamburg.

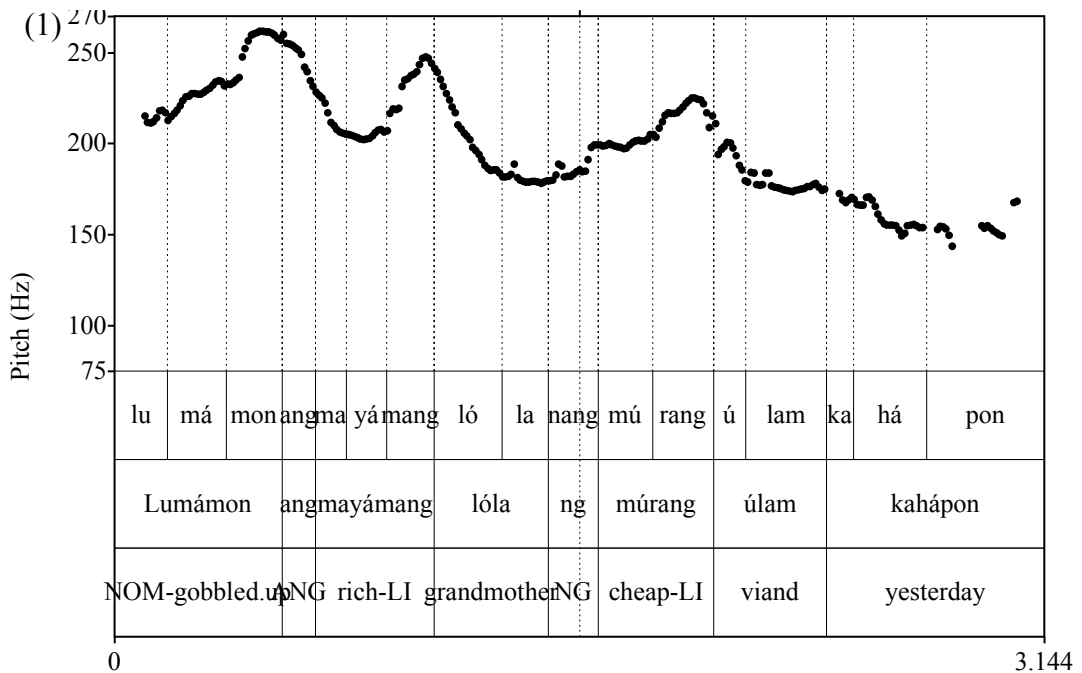
SOME NOTES ON TAGALOG PROSODY AND SCRAMBLING*

Norvin Richards
 MIT
 norvin@mit.edu

I sketch some basic facts of Tagalog prosody, including the position of pitch excursions in ordinary transitive sentences. We see that the object in VOS sentences is reliably higher pitched than the subject would be in VSO sentences.

1. Introduction

This paper will be concerned with facts like those in the Tagalog pitch track in (1):



‘The rich grandmother gobbled up the cheap viand¹ yesterday’²

* I am very grateful to Y-Chie S. Primo, Mervelina Saturno-Condon, Carmela Lao, Pearl Brault, and Karlyn Vito for everything they’ve taught me about their language. Thanks, too, to Joey Sabbagh, Kie Zuraw, Sun-Ah Jun, Henrison Hsieh, Lauren Clemens, Adam Albright, Edward Flemming, Nozomi Tanaka, Yuko Otsuka, two anonymous reviewers, and audiences at MIT and at ETI 3 and AFLA 22 at McGill University for discussion of the Tagalog facts. Any remaining mistakes are entirely my own responsibility.

I will first offer an account of the distribution of rises and falls in pitch in Tagalog statements like that in (1), mostly agreeing with the proposal of Sabbagh (this volume). I will then turn to the prosodic effects of scrambling.

2. Tagalog background

Tagalog is a verb-initial language, with scrambling in the postverbal field:

- (2) a. Lumunon ang ina ng mani
 NOM.swallowed ANG mother NG peanut
 'The mother swallowed a peanut'
- b. Lumunon ng mani ang ina
 NOM.swallowed NG peanut ANG mother
 'The mother swallowed a peanut'

I will use the conventions of Rackowski (2002) to gloss the Philippine-type 'voice' morphology. She analyzes this morphology as case agreement with a DP in the clause (for instance, the verb in (2) has NOM agreement, reflecting the fact that the thematic subject is picked out by the verb's morphology). The case markers on the DPs are ANG, indicating the DP with which the verb agrees; DAT, indicating Dative case; and NG, a default marker which appears on DPs not in the first two classes.

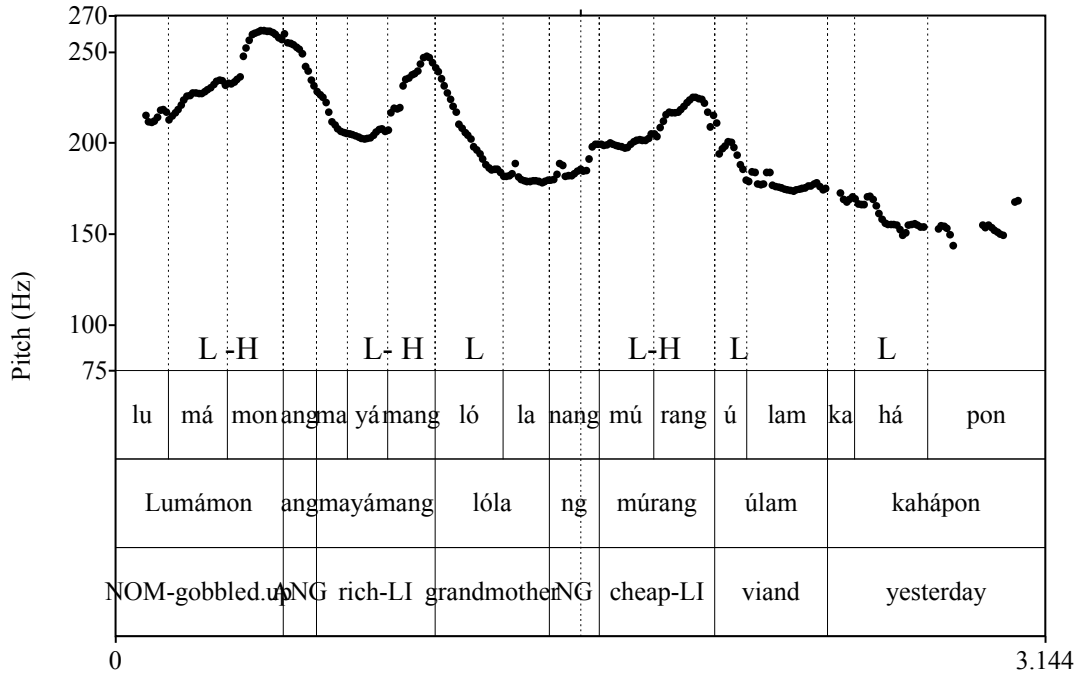
3. Tagalog prosody

As Sabbagh (this volume) also argues, the basic facts of Tagalog prosody follow from an explanation like that developed by Elfner (2012, to appear) for Connemara Irish. Tagalog prosody resembles its Irish counterpart, in that DPs generally begin with pitch rises and end with pitch falls, and the verb contains a pitch rise:

¹ *ulam* means 'the part of the meal which is not rice'; that is, traditionally, a standard meal consists of rice together with something else, and that something else is the *ulam*. 'Viand' is a standard translation for many Tagalog speakers, so I will use it here.

² In the segmental tier of pitch tracks, I will use Tagalog orthography, with two general exceptions: first, I will mark the position of stress, and second, the words *ng* 'NG' and *mga* 'plural' will be written out as pronounced, as 'nang' and 'manga', rather than being abbreviated as is standard in Tagalog. Tagalog orthography generally represents Tagalog pronunciation fairly accurately, though it does not represent glottal stops; orthographically vowel-initial words, for example, generally begin with glottal stops.

(3)



'The rich grandmother gobbled up the cheap viand yesterday'

The Tagalog example in (3) ends in an adverb, which means that Elfner's theory predicts, correctly, that both of the DPs in this example should contain a pitch rise (but that the adverb should lack one). Tagalog stress is typically either penultimate or word-final, unlike Irish stress which is generally word-initial; consequently, the Tagalog pitch rises are later in the word than they would be in Irish, a fact we can capture by associating these pitch movements, in both languages, with stressed syllables. I will refer interested readers to Elfner (2012, to appear) and Sabbagh (this volume) for further discussion of the Irish facts and of Elfner's theory.

Tagalog often, though not always, exhibits a 'pitch reset' between DPs. (3) offers a good example of this; the noun *lóla* 'grandmother', which ends the first DP, ends at a particularly low pitch, and the subsequent DP *ang múrang úlam* 'the cheap viand' begins by jumping to a slightly higher pitch. I will not discuss pitch reset further below, but careful readers may want to bear it in mind.

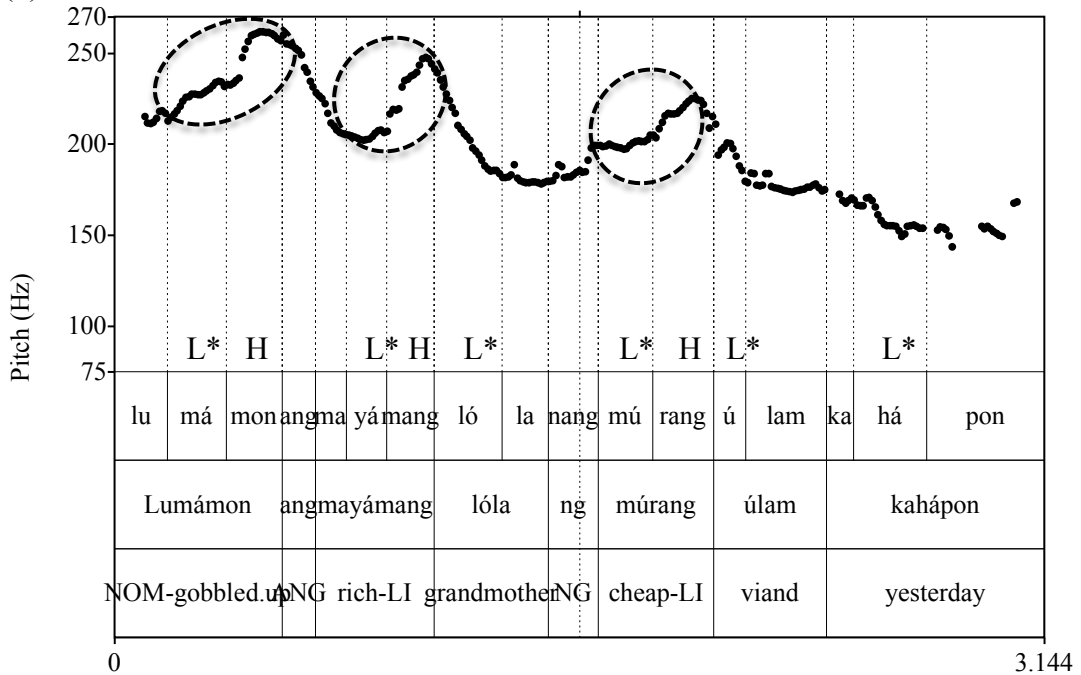
4. Rising and falling in Tagalog

In this section I will try to define a little more carefully the distribution and shape of pitch rises and falls in Tagalog. What we learn here will be useful in the next section, where we will turn to the interaction of prosody with scrambling. Section 4.1 will concentrate on the behavior of pitch rises; section 4.2 will turn to pitch falls, which exhibit some cross-speaker variation. Section 4.3 will summarize.

4.1. Pitch rises

By far the most common way to realize pitch rises in Tagalog involves a rise that begins at a stressed syllable. The pitch rise then typically peaks at the end of the word. Example (3), repeated below as (4), illustrates this general tendency:

(4)



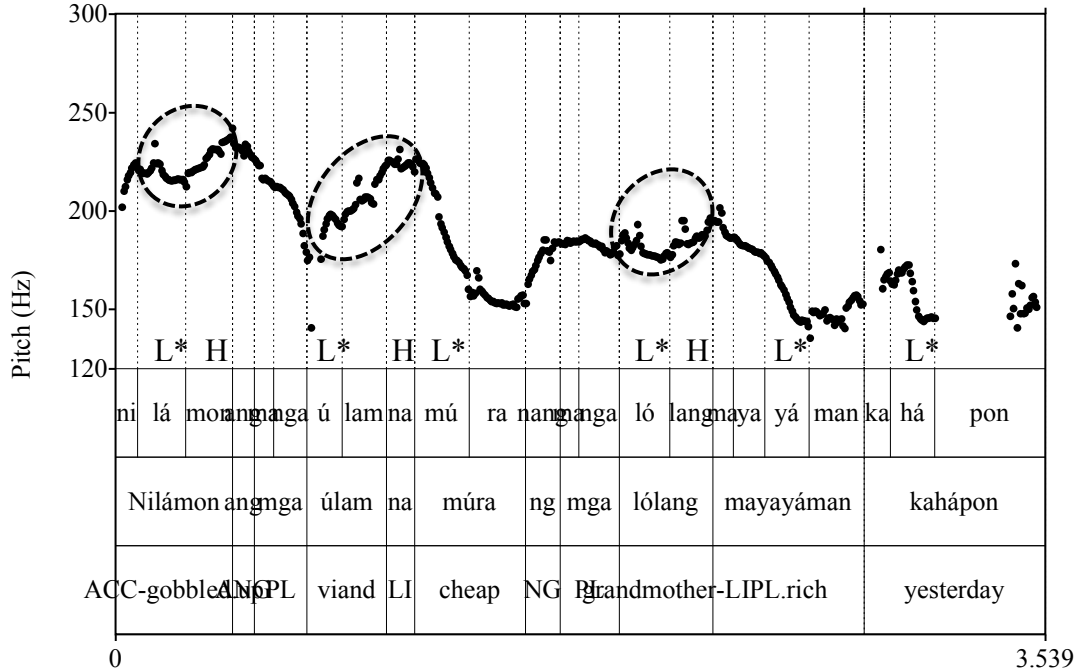
'The rich grandmother gobbled up the cheap viand yesterday'
[speaker A]

The example in (4) contains three pitch rises (circled), one on the verb and one on each of the adjectives modifying the two following DPs. Each pitch rise peaks in the syllable following a stressed syllable (which is, as it happens, the last syllable of the word in which the pitch peak begins). To make later discussion easier, I will begin labeling the speakers responsible for each utterance: the speaker responsible for example (4) will be known as speaker A.

The relevant notion of 'word' will have to be one in which a linker is considered to be part of the word preceding it, for purposes of pitch peak distribution³. Consider the examples in (5-6):

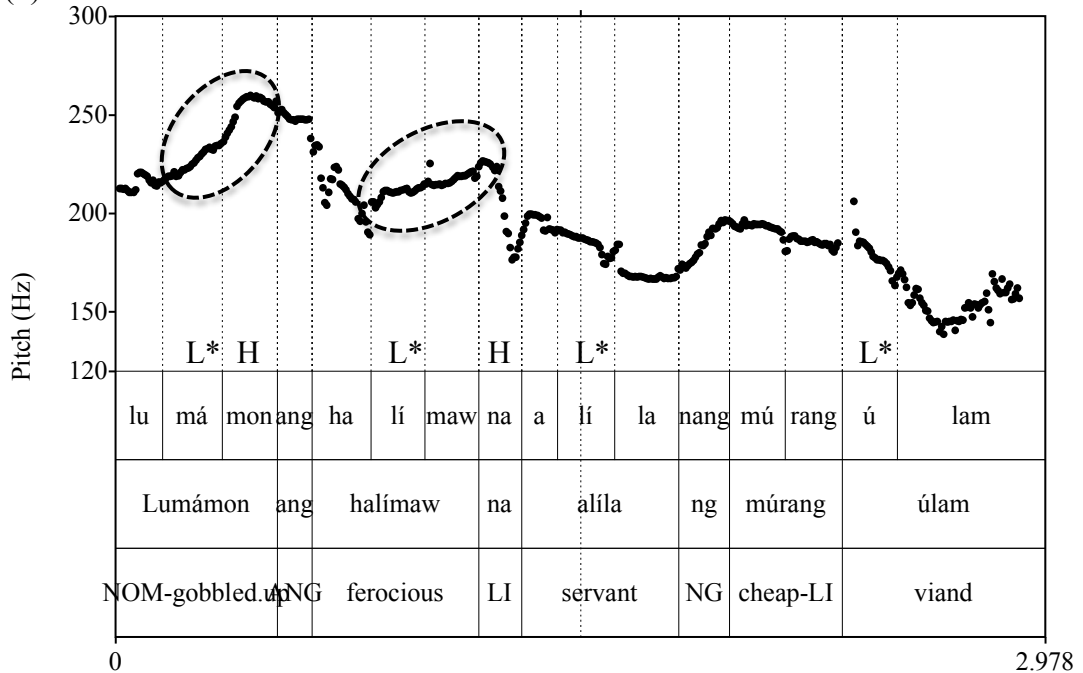
³ Lauren Clemens (p.c.) suggests, entirely plausibly, that the relevant notion of 'word' might be that of 'prosodic word'. I leave the issue open for future research.

(5)



'The rich grandmothers gobbled up the cheap viands yesterday' [speaker A]

(6)



'The ferocious servant gobbled up the cheap viand' [speaker A]

Here, again, the pitch rises are circled (and I am continuing to ignore the effects of pitch reset between DPs⁴). The pitch rise of interest, in both examples, is the second one, appearing at the beginning of the first DP (*úlam na múra* 'cheap viand' in (5), and *halímaw na alíla* 'ferocious servant' in (6)). Here the rise begins on the stressed syllable of the first word in the DP, and peaks on the linker *na*, which appears between the two words. The linker has two allomorphs, one a velar nasal attaching to the preceding word, and the other a free-standing syllable *na*. Apparently both allomorphs must be regarded as part of the preceding word, for purposes of determining the placement of pitch peaks.

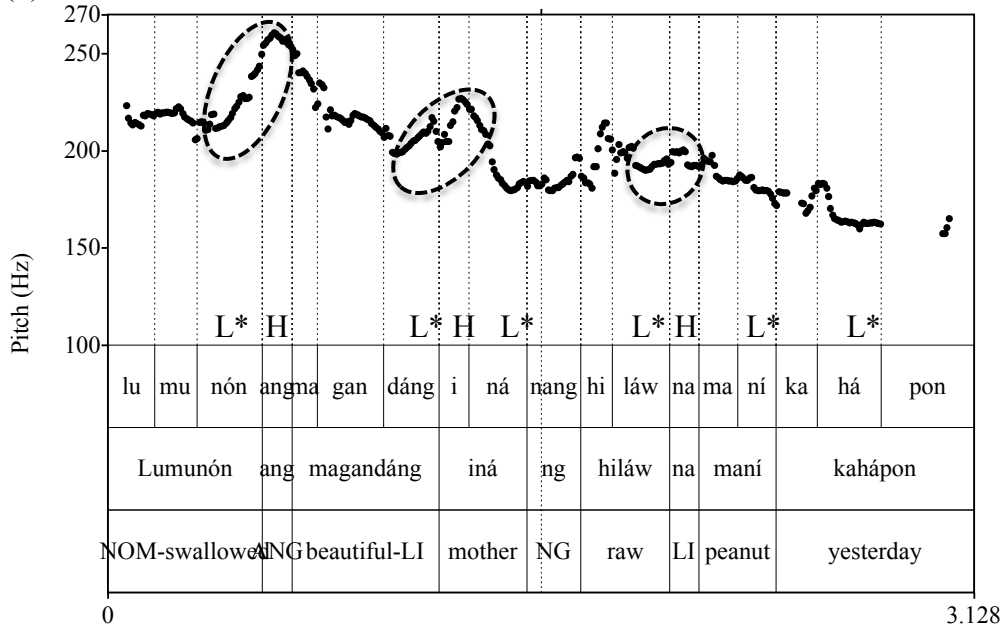
In this regard, the linker differs from other functional material. Consider, in (5-6), the placement of the first pitch peak, associated with the verb. This appears at the end of the verb itself, and is followed by a fall through functional morphemes beginning the DP (a case-marked determiner *ang*, and, in (6), the bisyllabic plural morpheme *mga* /maŋa/). These morphemes do not participate in the verb's pitch rise, in the way that the linker participates in the pitch rise in the first DP.

We also learn from examples like (5-6) that pitch rises do not simply peak on the syllable after they begin (as they did in example (4) above). In both of (5-6), the second pitch rise of the sentence begins at the penultimate syllable of a word, and passes through the word's final syllable to peak at the linker.

Tagalog main word stress is generally either penultimate or final, which will make it difficult to observe pitch peaks any further from stressed syllables than those in (5-6). Pitch peaks associated with final stress typically appear on the immediately following syllable, or sometimes at the transition between the stressed syllable and the following one:

⁴ Note that pitch reset is clearly not simply another instance of a pitch rise, of the same type as the ones circled. In (11), for example, the second DP *ang mga lolang mayayaman* 'the rich grandmothers' begins with a higher pitch than the last syllable of the immediately preceding adjective *mura* 'cheap'. But this rise in pitch cannot be triggered by the stressed syllable of *mura*; if it were, then we would expect the pitch rise to begin on that stressed syllable and continue through the final syllable of *mura*. Compare the pitch rise on the preceding noun *ulam* 'viand', which does begin with the initial syllable of the noun. "Pitch reset" appears to be a distinct phenomenon.

(7)



'The beautiful mother swallowed a raw peanut yesterday' [speaker A]

The content words in (7) generally have final stress, and their pitch rises begin on their final syllables, and peak on the first syllable of the following word.

In general, then, pitch rises in Tagalog appear to involve an L* associated with the stressed syllable, followed by an H which appears at the end of the word (and linkers are associated tightly enough with the preceding word to count as part of it for this purpose). Just when the stressed syllable is itself at the end of the word, the associated H will often appear at the beginning of the next word.

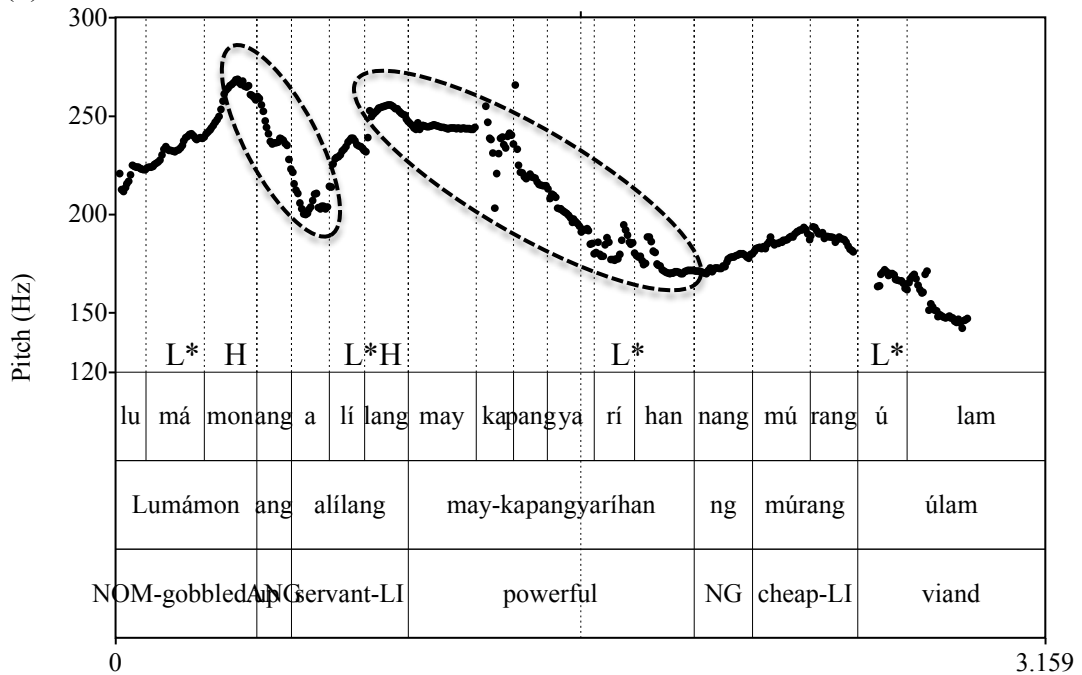
4.2. Pitch falls

Pitch falls are realized differently by different speakers; I will suggest that for some speakers, a pitch fall is a simple L* on a stressed syllable, while for others, the L* is preceded by an H target.

For speakers like speaker A, pitch falls start at the pitch peak of the preceding rise, and end at a stressed syllable. The pitch drop is often steepest on the stressed syllable itself, and comparatively gradual on preceding syllables. Material following the stressed syllable is typically either flat or continues a more gradual fall than the one on the stressed syllable. In this pattern, there is no evidence for a H target associated with the pitch fall; it is simply an L*.

Consider example (8):

(8)

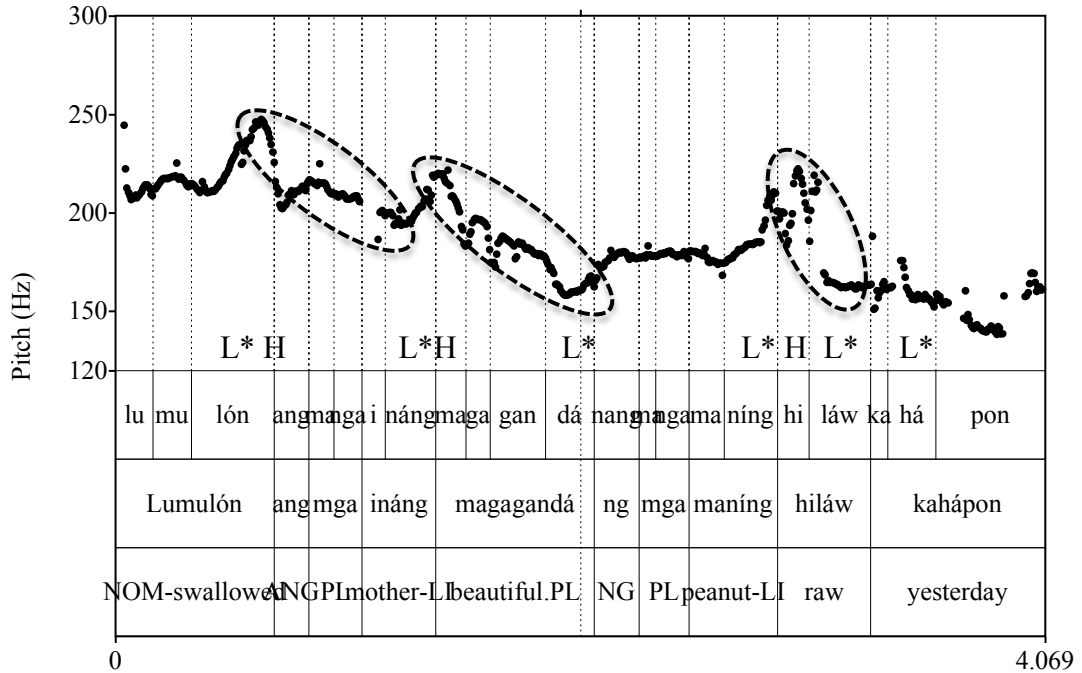


'The powerful servant gobbled up the cheap viand' [speaker A]

In (8), two pitch falls are circled. The first DP has a quick pitch rise which begins on the penultimate syllable of *alilang* 'servant-LINKER', and ends on its final syllable. The subsequent fall spans six syllables, running from the pitch peak at the final syllable of the noun to the stressed syllable of the following adjective *may-kapangyarihan* 'powerful'.

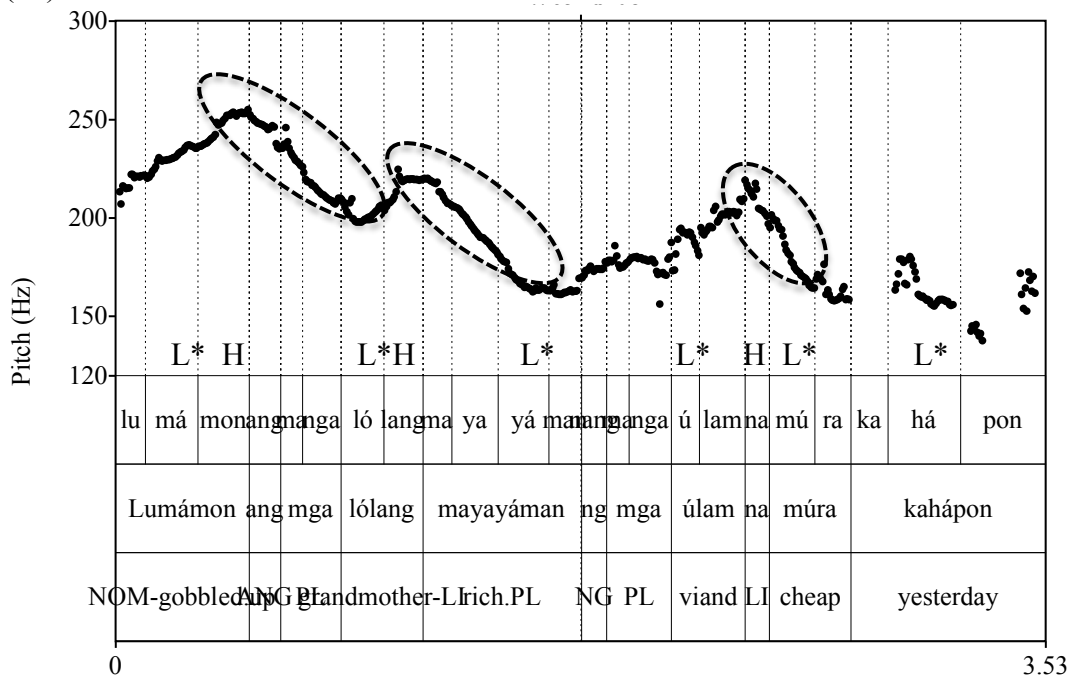
In (9-10), we see some additional evidence that the L involved in pitch falls is associated with the position of stress:

(9)



'The beautiful mothers swallowed raw peanuts yesterday' [speaker A]

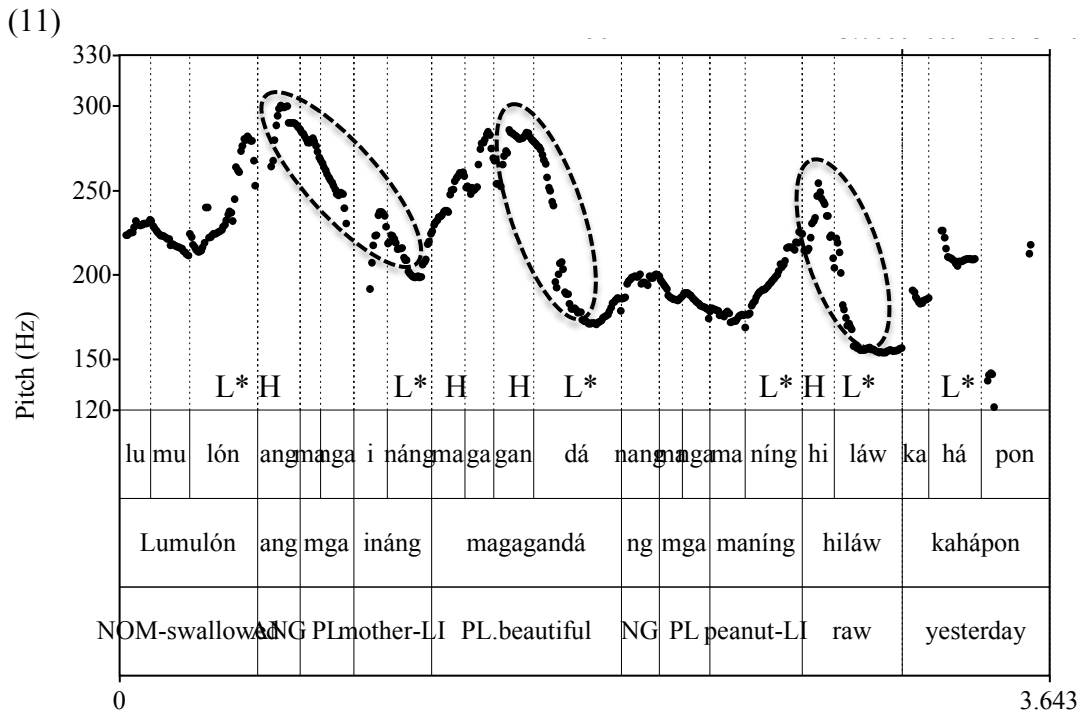
(10)



'The rich grandmothers gobbled up the cheap viands yesterday' [speaker A]

Consider, in particular, the properties of the second pitch fall, the one at the end of the first DP of each of these sentences. In both cases, the pitch fall ends on a stressed syllable; stress is final on the adjective *magagandá* 'beautiful.PL' in (9), and penultimate on the adjective *mayayáman* 'rich.PL' in (10). The examples illustrate the fact that pitch falls end at stressed syllables (rather than, for example, at the ends of words or phrases).

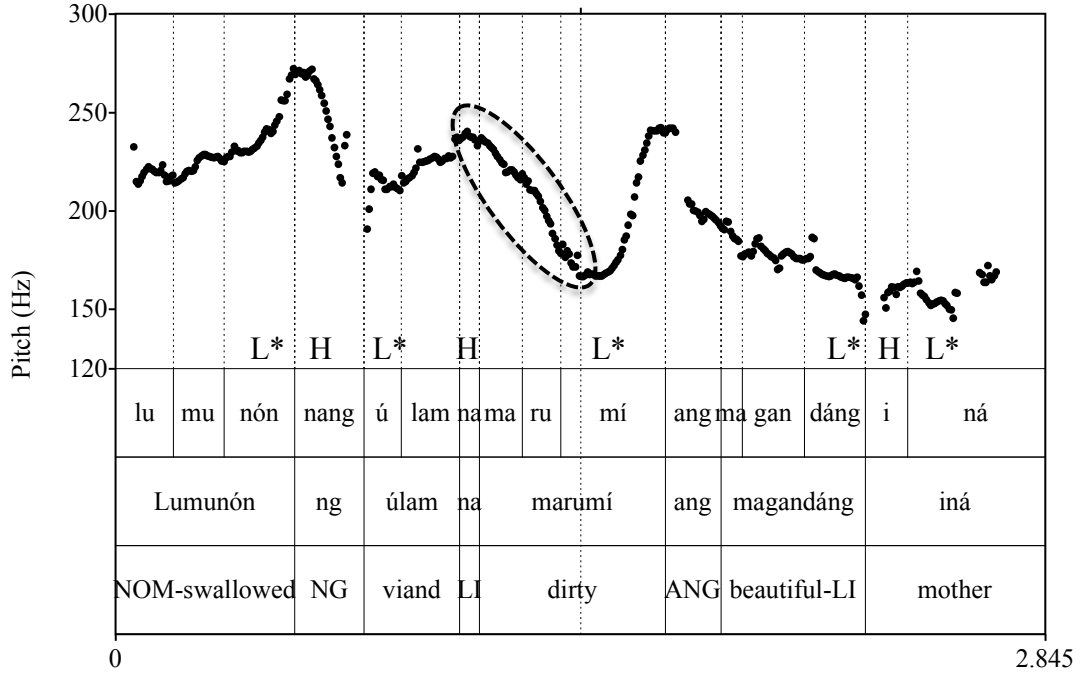
There is another well-attested way of implementing pitch falls, which I will suggest represents association of phrase-final stress with a H L* accent, rather than the L* used by the speaker above. Consider (11), a pitch track for the sentence in (9) above, as uttered by a different speaker, referred to hereafter as speaker B:



'The beautiful mothers swallowed raw peanuts yesterday' [speaker B]

Three pitch falls are circled in (11), and the one of interest to us is the second one, on the DP *ang mga ináng magagandá* 'the beautiful mothers'. For speaker B, this fall is quite steep, contained almost entirely in the final syllable of the adjective. The steepness of the fall is puzzling on the model developed so far, which would lead us to expect a pitch peak just after the first stressed syllable of the DP, with stress falling gradually throughout the adjective, to its stressed final syllable. This was, in fact, the pattern that we saw in (9) above, for speaker A. Speaker B in (11), however, has a different pattern. The difference is not simply a quirk of these particular examples; another pair of examples with the same property follows:

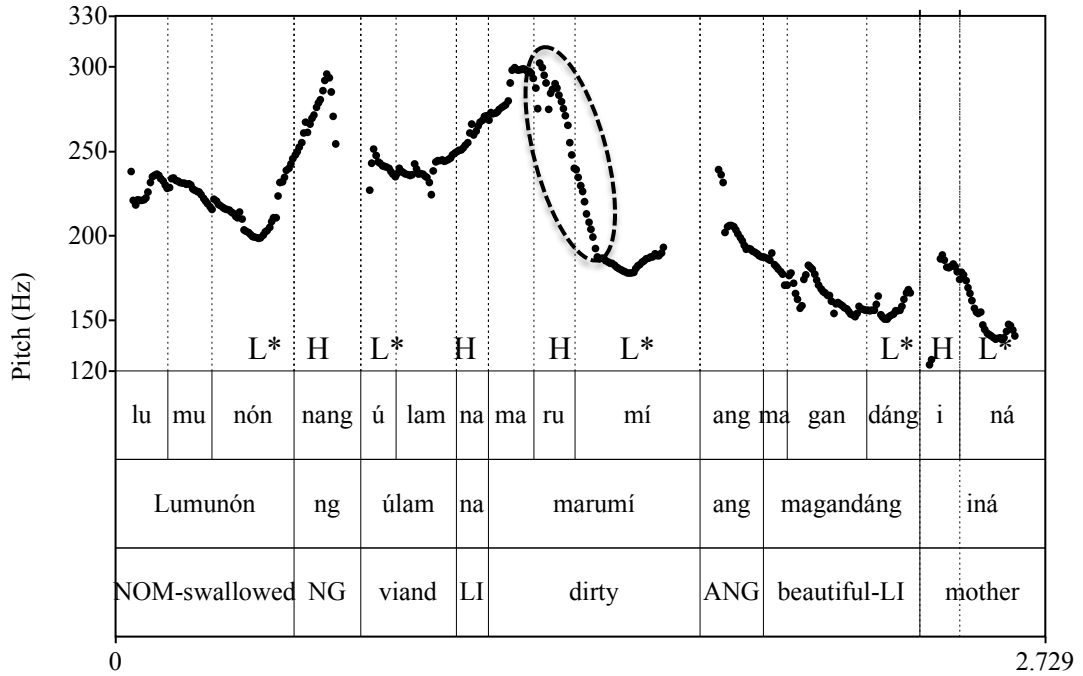
(12) a.



'The beautiful mother swallowed a dirty viand'

[speaker A]

b.



'The beautiful mother swallowed a dirty viand'

[speaker B]

Here, again, the pitch fall in the first DP (circled) is steeper for speaker B than for speaker A. For speaker A, pitch begins falling around the beginning of the adjective, after the pitch peak on the linker induced by the stressed syllable in the noun. For speaker B, the pitch fall starts in the penultimate syllable of the adjective, just before the stressed final syllable.

The explanation for the difference should not be about different realizations of pitch rises more generally. We cannot, for example, say that speaker B simply has a different implementation for pitch rises than speaker A, one in which pitch rises until forced to fall by a stressed syllable. Such a description of pitch rises could account for the behavior of the second pitch peak in (11), but not for the first pitch peak. The pitch peak on the verb is just where it should be in (11), on the syllable after the stressed final syllable of the verb, and pitch after this peak drops comparatively gradually to the stressed syllable on the noun *ináng*, just as it would for speaker A. And, again, this is a general feature of this speaker's pronunciation; speakers A and B treat verbs in the same way, but differ in their treatment of DPs.

The best account of the difference between speakers, then, would need to account for a difference between two kinds of cases: one case in which a pitch rise is followed by another pitch rise (in this case, the rise on the verb, followed by the rise at the beginning of the DP), and another case in which a pitch rise is followed by a pitch fall (the case of the two pitch excursions associated with DP). For speaker A, these two cases are identical; they involve a rise followed by an L* tone, which is either the beginning of a following rise (L* H) or a phrase-final fall (L*). For speaker B, apparently, the situation is different. We can capture the difference between speakers by representing phrase-final falls for speaker B as H L*. These falls therefore begin, not with an L*, but with an H, and consequently, the L* of the DP-final fall is always just preceded by an H; the fall is therefore not as gradual as it would be for speaker A.

4.3. Summary

The general principles sketched above are summarized in (13-14):

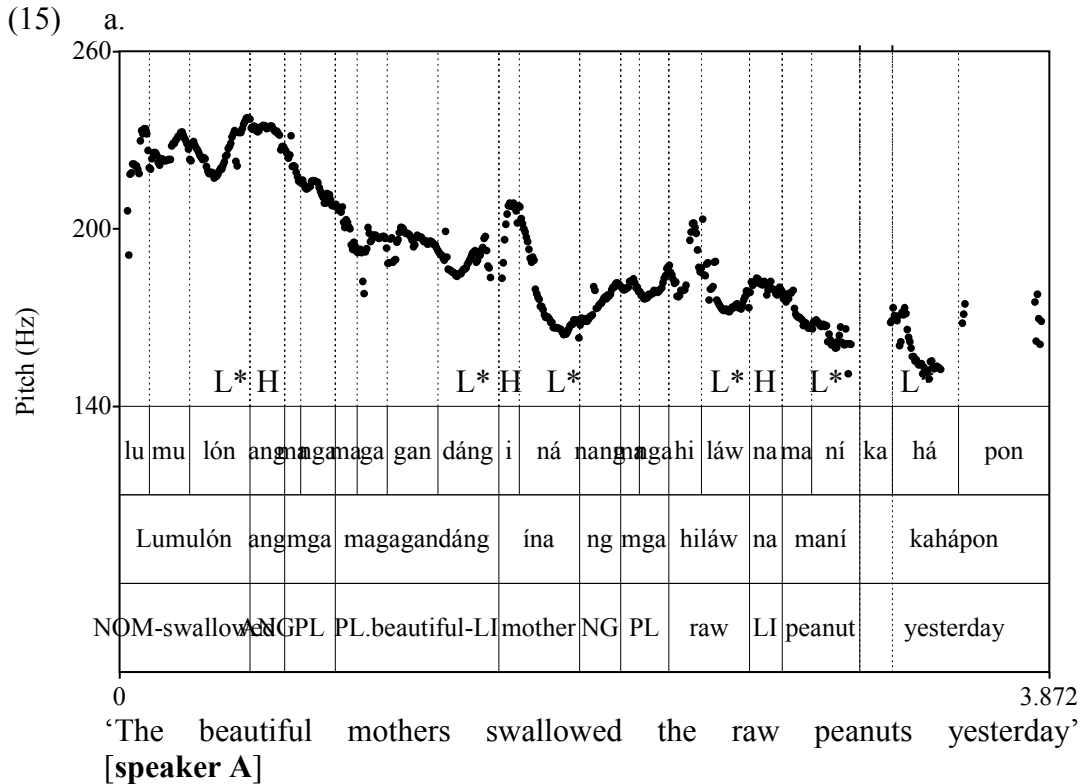
- (13) A pitch rise (L* H) begins at a stressed syllable, and peaks at the end of the word (with linkers treated as part of the preceding word). If the stressed syllable is final, the peak is sometimes on the first syllable of the following word.
- (14) Pitch falls have at least two attested realizations.
 - a. For some speakers, a pitch fall begins at a pitch peak and proceeds to a stressed syllable.
 - b. For other speakers, a pitch fall begins at the syllable before a stressed syllable, and ends on the stressed syllable. (H L*)

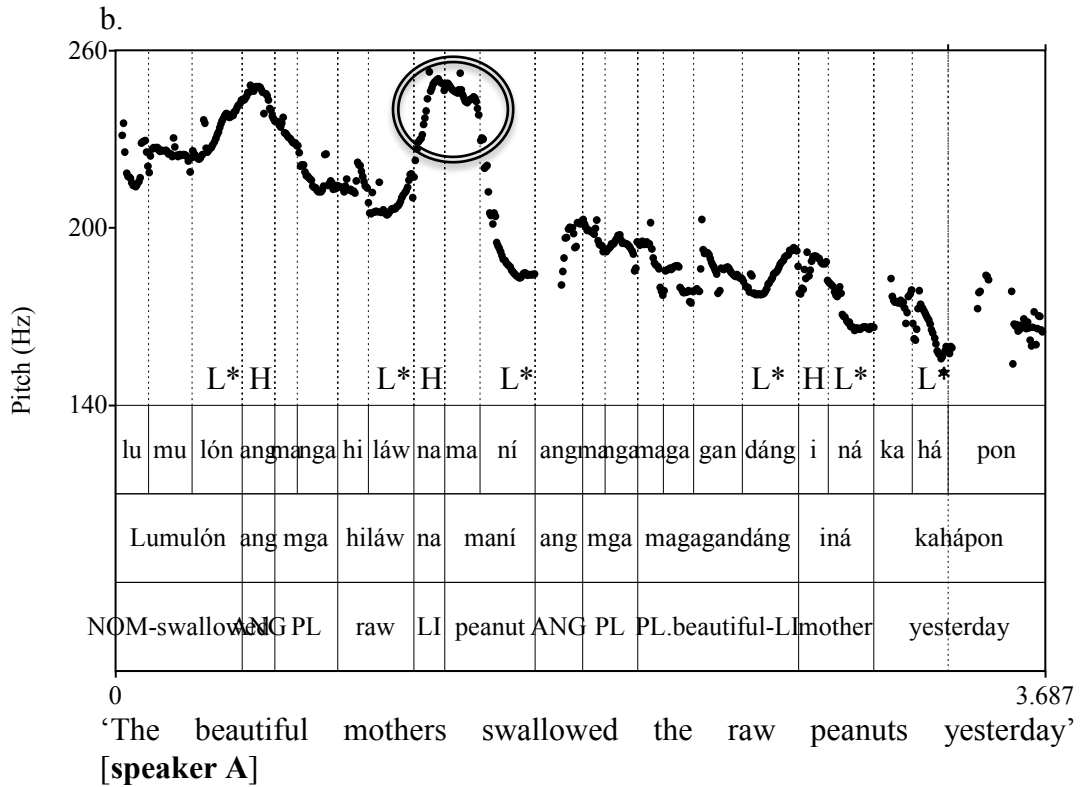
We have seen one systematic difference between speakers A and B, which I proposed had to do with a difference in the representation of phrase-final falls; for

speaker A, phrase-final falls are triggered by an L*, while for speaker B they involve a H L* sequence. This does not exhaust the differences between speakers in the sample, but I will stop here for reasons for space.

5. Prosody and word order

We now turn to an interaction between scrambling and the relative height of certain pitch peaks. Consider the examples in (15):





(15a) has VSO word order, and (15b) VOS. The circled peak in (15b) emphasizes a prosodic difference between the two examples; when the object precedes the subject, its pitch peak is very high, much higher than the corresponding peak on the subject in (15a). As we will see, this pair of examples is representative.

Four native speakers of Tagalog, all women from the Manila area now living in the Boston area, were presented with 16 transitive clauses, in which both DPs contained both a noun and a modifying adjective, together with 50 fillers. The 16 clauses varied along four binary parameters. Their content words had either consistently penultimate or consistently final stress; the 'voice' morphology on their verbs was either NOM or ACC; the adjectives modifying their nouns were either prenominal or postnominal; and their word order was either VSO or VOS. Two of the relevant examples, differing in placement of stress, were as follows:

- (16)a. Lumulón ang magandáng iná ng hilów na maní kahápon
NOM.swallowed ANG beautiful-LI mother NG raw LI peanut yesterday
'The beautiful mother swallowed a raw peanut yesterday'
- b. Lumámon ang mayámang lóla ng múrang úlam kahápon
NOM.gobbled ANG rich-LI g.ma NG cheap-LI viand yesterday
'The rich grandmother gobbled up the cheap viand yesterday'

The speakers were asked to read the sentences at ordinary conversational speed, imagining that they represented the beginning of a conversation, as an answer to a question like 'What happened?'. They were encouraged to repeat any sentences that they felt they had pronounced oddly. They were recorded in a quiet room, using a Marantz PMD 670.

I used Praat to create pitch tracks of the resulting recordings, and to find the highest pitch points in the verb and in each of the DPs of the sentence (relying mainly on Praat's automatic function for finding highest pitches, with occasional corrections by hand for obviously disturbed pitch tracks). As the preceding sections have hopefully made clear, the highest point in the DP could have appeared in any of a number of places, with the most common position being one or two syllables after the first stressed syllable in the DP. I recorded the highest position anywhere in the DP, without confining myself to this window.

For speaker A, the average results were the following, broken down by word order (each row in the following table represents the average of 8 utterances with the relevant word order):

(17)

	Verb peak	first DP	second DP	[speaker A]
VSO	246.9	219.6	206.4	
VOS	244.3	239.7	200.6	

The data in (17) illustrate the generality of the effect illustrated in the examples in (15); on average, objects have a significantly higher pitch than subjects in immediate postverbal position ($p=0.0116$), and the difference in peak height between the peaks of the two DPs is significantly greater in VOS order ($p=0.0006$). In (18), the same data are broken down by 'voice' morphology:

(18)

	Verb peak	first DP	second DP	[speaker A]
NOM, VSO	248.1	220	210.9	
NOM, VOS	241.6	237	197.6	
ACC, VSO	240	214.3	203.7	
ACC, VOS	251.5	242.8	204.3	

A full discussion of the effects of 'voice' is beyond the scope of this paper⁵, but the data in (18) suggest that the facts reported in (17) are consistent across the two 'voices'; note that the first DP is higher in the second and fourth lines (when it is an object) than in the first and third (when it is a subject). The relevant notion of 'subject', for the generalization about pitch height, is apparently about the thematic subject, rather than the DP picked out by the verb's 'voice' morphology.

Speakers B, C, and D, like speaker A, have a greater distance in pitch between the pitch peaks of the two DPs in VOS order than in VSO order:

⁵ For more discussion of the relevant effects, see Hsieh and Travis (2014) and Hsieh (2015).

(19)

	Verb peak	first DP	second DP
VSO	317	279.8	260
VOS	300	292.8	249.6

 [speaker B]

(20)

	Verb peak	first DP	second DP
VSO	279.1	226.6	214.2
VOS	288.8	235.3	209.7

 [speaker C]

(21)

	Verb peak	first DP	second DP
VSO	270.9	248.7	233
VOS	278.7	273	220

 [speaker D]

As the tables in (19-21) show, all of these speakers exhibit a greater distance in pitch between the first and second DP in VOS word order than in VSO. For speaker A, this distance was mainly created by boosting the pitch of the scrambled object; in (19-21), we can see that other speakers both boost the pitch of the object and suppress the pitch of the subject. Pooling all of the data for the four speakers, the greater distance in pitch between the two DPs in VOS order is statistically significant ($p=0.0056$).

These differences in pitch might be the expression in Tagalog of a cross-linguistic tendency for internal arguments to receive greater stress than external arguments, a phenomenon known as 'nuclear stress' (Chomsky and Halle 1968, Halle and Vergnaud 1987, Cinque 1993, Zubizarreta 1998, Arregi 2002, Kahnemuyipour 2009, and references cited there), represented here with small capitals:

(22) The woman bought a BOOK.

The most natural pronunciation of (22), in wide-focus contexts, involves main sentential stress on *book*, and not (for example) on the subject. Similarly, in Tagalog, in the wide-focus contexts in which the data discussed above were gathered, an object will have higher pitch than a subject in the same position.

We saw above that the special prosodic status of objects in Tagalog persists even when basic word order is altered by scrambling. Tagalog scrambling, on this view, would have to be like the kinds of English A-bar movement Bresnan (1972) discusses, and unlike, for example, Basque scrambling (Arregi 2002) or Persian object shift (Kahnemuyipour 2009), both of which have been argued to alter the distribution of nuclear stress.

This view of the Tagalog prosodic data has the virtue of placing Tagalog in a fairly familiar context; Tagalog is yet another language, apparently like every other language in which the facts have been carefully investigated, in which nuclear stress appears on objects rather than on subjects. Much of the Tagalog literature is focused on defining the notion of 'subject' for this language, in a way

which makes reference to its infamous 'voice' system. The nuclear stress facts, if that is what they are, seem to make reference to an entirely conventional argument structure, in which 'subjects', for the relevant purposes, are simply the arguments which receive the highest theta-role in the clause.

6. Conclusions

This paper has been an attempt to describe the intonation of a language in which intonation has so far been fairly scantily described. After proposing, in agreement with Sabbagh (this volume), that the distribution of rises and falls in Tagalog may be captured by Elfner's (2012) account of the prosody of Irish, I went on to try to define with some precision the distribution of pitch rises and falls in Tagalog (noting certain instances of cross-speaker variation). I also discussed an intriguing difference between internal and external arguments in Tagalog; internal arguments have higher pitch than subjects would in the same position, a difference I proposed to describe as an instance of the effects of nuclear stress. Much work on Tagalog intonation remains to be done, some of which I hope to be able to do in future.

References

- Arregi, Karlos. 2002. Focus on Basque movements. Doctoral dissertation, MIT.
- Bresnan, Joan. 1971. Sentence stress and syntactic transformations. *Language* 47.257-281.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24:239-297.
- Elfner, Emily. 2012. Syntax-prosody interactions in Irish. Doctoral dissertation, University of Massachusetts, Amherst.
- Elfner, Emily. To appear. Recursion in prosodic phrasing: evidence from Connemara Irish. *Natural Language and Linguistic Theory*.
- Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10:375-414.
- Halle, Morris, and Jean-Roger Vergnaud. 1987. *An essay on stress*. Cambridge, MA: MIT Press.
- Hsieh, Henrison. 2015. Phrase structure via prosody in Tagalog. Ms., McGill University, Montreal, Canada.
- Hsieh, Henrison, and Lisa Travis. 2014. Tagalog phrase structure and preliminary prosodic results. Paper presented at Exploring the Interfaces 3: Prosody and Information Structure.
- Kahnemuyipour, Arsalan. 2009. *The syntax of sentential stress*. Oxford University Press.

The Proceedings of AFLA 22

- Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.
- Rackowski, Andrea. 2002. The structure of Tagalog: specificity, voice, and the distribution of arguments. Doctoral dissertation, MIT.
- Sabbagh, Joey. This volume.
- Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. Cambridge, Ma: MIT Press.

GENERAL NUMBER AND SEMANTIC INCORPORATION IN TAGALOG*

Nova Starr
University of Ottawa
novastarr@uottawa.ca

This paper expands on Tagalog general number (GN) by identifying a much more specific context for where GN can and cannot occur; unmodified *ng* objects have GN. What's more is that word order has no effect on this. GN is often considered a semantic indicator of noun incorporation (NI) (Carlson 2006, Dayal 2011, Farkas and de Swart 2003, Mithun 1984) however my analysis finds that while Tagalog GN shares the semantic properties of NI, it does not share the structural properties. In traditional NI, the incorporated nominal (IN) maintains a close relationship to the verb (Massam 2001, Johns 2007). Therefore, I describe Tagalog as a language with semantic incorporation. The semantic framework for my analysis applies Chung and Ladusaw (2004)'s *Restriction and Saturation* to the Tagalog data. This framework treats the differences between sentences with incorporation and sentences without it as differences in modes of composition. In addition to this work, I draw parallels between previous syntactic analyses of specificity and my description of voice dependent general number.

1. Introduction

The voice system of Tagalog has been a hot topic for Austronesianists, in particular, its semantic interpretation. Each voice has several semantic consequences and in order to get any closer to determining which voice is the least marked, examining these semantic consequences is crucial.

The aim of the present paper is twofold. First, I identify an additional voice dependent semantic feature: general number (GN). General number is described as a nominal form that is interpreted as entailing 'one or more x' (Corbett 2000). Second, I give an analysis correlating GN in Tagalog to noun incorporation (NI). This builds

*Major thanks go to my supervisors Dr. Robert Truswell and Dr. Éric Mathieu. Éric gave a lot of time and constructive feedback through out the process. His expertise in incorporation has been an immense asset to this project. Even all the way from Scotland and through typhoons and power outtages while I was in the Philippines, Rob has guided me through this whole process from day one with countless Skype calls and emails. He helped me look at this project from many different angles. I would also like to thank my consultants, in particular Henrison Hsieh for his patience with my constant questions. And last but not least, I am so grateful for the constructive feedback I received from several experts at AFLA as well as at the CLA. It has been an incredible learning experience. Thank you! Maraming salamat sa inyo!

on the established generalization that GN is characteristic of NI (Carlson 2006, Dayal 1999, 2011, Farkas and de Swart 2003). My analysis finds that while GN in Tagalog shares the semantic features of NI, it does not share the syntactic features.

The present paper is organized as follows: I describe GN in detail and consider its links with incorporation in section 2. In section 3, I present the structural framework I adopt with respect to voice and word order. In section 4, I mention some analyses of incorporation but more thoroughly detail the mechanics of Chung and Ladusaw (2004). Next, in section 5, I borrow Chung and Ladusaw (2004)'s analysis and apply it to Tagalog. Finally, I conclude in section 6.

2. Data and determining *ng* and *ang*

In this section, I briefly review Paul et al. (In Press)'s semantic description of *ng* and *ang* as well as introduce the restricted context in which Tagalog nominals have GN.

2.1. Definiteness of *ng* and *ang* by Paul et al. (In Press)

Earlier descriptions have labelled *ang* as definite and *ng* as indefinite (Schachter and Otanes 1972) however more recent research, in particular, Paul et al. (In Press) have presented data using semantic tests to demonstrate that while *ang* and *ng* do share some properties with definites and indefinites respectively, they do not share all of the properties. This is evident when contrasting the unmodified particle with its counterpart that is modified with *isang* which means one. The semantic diagnostics used aim to describe *ang* and *ng* with respect to referentiality, scope, and sluicing.

2.1.1. *ng*

Semantically, not all *ng* arguments are the same. Some speakers find that subjects, whether modified by *isang* or not, must take wide scope whereas objects may have either wide or narrow scope. With respect to referentiality and sluices, *ng* arguments do not distinguish between subjects or objects. Unmodified *ng* arguments may denote either a novel or familiar referent while *ng isang* arguments must denote a novel referent. Both *ng* and *ng isang* arguments can serve as correlates in sluices. Paul et al. (In Press) concludes that *ng* arguments share most of the characteristics of indefinites but not all. Crucially, *ng* arguments can be interpreted as familiar (in addition to novel), therefore resembling definites.

2.1.2. *ang*

Paul et al. (In Press) finds that the definiteness of *ang* is even more unclear than that of *ng* arguments. For all three diagnostics, *ang* and *ang isang* arguments differ. The former denote familiar referents while the latter denote novel referents (Paul et al. In Press:13). Furthermore, the former are scopeless (or take wide scope) and the latter can take variable scope. Finally, with respect to sluices, unmodified *ang* arguments do

not serve as correlates to *wh*-phrases but *ang isang* arguments do. So while unmodified *ang* arguments appear definite, *ang isang* arguments appear indefinite.

2.1.3. Remarks

Where *isang* is present, *ang* and *ng* have no interpretational differences. However this changes when *isang* is absent. Alone, *ang* appears to denote definiteness and *ng* appears to denote indefiniteness. This suggests that there may be external forces that control the interpretation of *ang* and *ng*.

To conclude, table 1 summarizes the semantic findings of both particles, *ang* and *ng*, as well as their modified counterparts, *ang isang* and *ng isang*. The analysis that follows in section 5 treats *ang* as a definite determiner. This treatment should be considered a simplification for the purpose of making space for my analysis of *ng* arguments.

	<i>ang</i>	<i>ang isang</i>	<i>ng</i>	<i>ng isang</i>
Referentiality	familiar	novel	novel/familiar	novel
Scope	wide (scpless)	wide/narrow	wide/narrow ¹	wide/narrow ²
Sluice	not correlate	correlate	correlate	correlate

Table 1: Summary of semantic tests for *ang* and *ng* (Paul et al. In Press)

2.2. General number

The plural particle, *mga*, signifies that its associated nominal is plural. Observe the following two examples, (1) and (2), where the object can only be interpreted as plural. Next, observe the morphologically plural subjects in (3) and (4).

- (1) B⟨um⟩ili ng mga libro ang bata. (2) B⟨in⟩ili ang mga libro ng bata.
av.buy ng pl book ang child pv.buy ang pl book ng child
 ‘The child bought books.’ ‘The/A child bought the books.’
- (3) B⟨um⟩ili ng libro ang mga bata. (4) B⟨in⟩ili ang libro ng mga bata.
av.buy ng book ang pl child pv.buy ang book ang pl child
 ‘The children bought one or more books.’ ‘Children bought the book.’

Number is not always morphologically marked. In some contexts, however, nominals without the plural particle, *mga*, can still indicate a plural reading. The

¹Some speakers only accept wide scope interpretations for *ngA (isang)* arguments.

²See footnote 1.

object in (3) is one example of this. This object is interpreted as ‘one or more books’. Conversely, the object in (4) can only denote a single item, ‘the book’.

General number (GN) is described as a nominal form that is interpreted as entailing ‘one or more *x*’ (Corbett 2000). The English sentence, (5), does not signify the number of berries picked. It could be one or more.³ It also denotes a type of picking activity and directs focus away from the berries that may have been picked. In contrast, (6) and (7) do not have GN. The meaning of the former is that a single berry was picked and the latter means that several berries were picked.

(5) Susan went berry-picking. (6) Susan picked a berry.

(7) Susan picked berries.

Corbett (2000) describes GN in Tagalog such that GN nominals and singular nominals share the same morphology. What has not been previously noted in the literature is the very specific environment where GN occurs. GN in Tagalog is conditioned by argument structure, voice, and bareness of the argument in question.

First, subjects never have GN, only objects do. In (1), there is exactly one child. Second, only objects in agent voice have GN. Compare the object in the agent voice example (3) with the object in the patient voice example (4). In (4), there is exactly one book that the child bought.

There is another restriction on Tagalog GN. Arguments with GN must be unmodified. Modified nominals, on the other hand, do not have GN. The modified object in (8) has a singular reading. The plural particle *mga* in (9) forces a plural interpretation of *libro* ‘book’.

(8) B(um)ili ng pula=ng libro ang bata.
av.buy ng red=lk book ang child
‘The child bought a red book.’

(9) B(um)ili ng mga libro ang bata.
av.buy ng pl book ang child
‘The child bought books.’

The above data leads to the inference that GN is a general property of unmodified direct objects that are *ng* marked. The ditransitive data presented below confirms this. The following ditransitive constructions illustrates this in patient voice, agent voice, and benefactive voice. The direct object in (10) is *ang* marked and has the interpretation of there being only one mango. Conversely, the direct objects in (11) and (12) are *ng* marked and have GN interpretations.

(10) B(in)igay ang mangga ng tatay sa bata.
pv.give ang mango ng father to child

³Technically, it could also mean that she went berry-picking but didn’t find any to pick. This detail isn’t discussed in (Corbett 2000) but I speculate that this may be a language-specific feature.

‘The father gave the mango to a/the child.’

- (11) Nag-bigay ng mangga ang tatay sa bata.
av-give ng mango ang father to child
‘The father gave one or more mangoes to a/the child.’
- (12) B(in)igyan ng mangga ng tatay ang bata.
bv.give ng mango ng father ang child
‘The father gave one or more mangoes to the child.’

2.3. Looks like incorporation, but is it?

GN is a hallmark property of noun incorporation (NI) (Carlson 2006, Dayal 1999, 2011, Farkas and de Swart 2003). In addition to GN, Tagalog displays other cross-linguistic properties of NI. Thus, it is reasonable to conclude that Tagalog is an incorporating language. Before expanding on this, Massam (2009:1078) considers various definitions of noun incorporation and provides a broad one:

“...a grammatical construction where a nominal that would canonically be expressed as an independent argument or adjunct is instead in some way incorporated into the verbal element of the sentence, forming part of the predicate.”

There are two central characteristics of Tagalog GN that align cross-linguistically with NI. First, most NI languages only incorporate the object (Baker 1988:81). Some languages that only incorporate objects are Mohawk (Postal 1962), Southern Tiwa (Allen et al. 1984), and Niuean (Seiter 1980). Second, many incorporating languages only incorporate bare nouns (Mithun 1984, Baker 1988). This means that incorporated nominals (INs) often lack case, determiners, and/or modifiers (Baker 1988). In Tagalog, *ng* argument is considered bare despite the presence of *ng*. Thus a simple generalization of the data given is that agent voice sentences have NI and patient voice sentences do not.

One of the main tenets of NI is that the IN must be morphologically attached or in close association with the verb (Johns 2007:537). Up to this point, I have only presented sentences with verb-object-subject (VOS) word order showing the verb as being immediately followed by the incorporated object. Tagalog has free post-verbal word order such that verb-subject-object (VSO) word order is also grammatical.

The most interesting thing about this flexibility in word order is that it makes no semantic difference whether a sentence has VOS or VSO word order. Scope, specificity, and also GN are the same regardless of the distance between the verb and the object. If a sentence with VOS word order has a GN interpretation, this interpretation is maintained in a similar sentence with VSO word order. For example, the agent voice sentence (13) is semantically identical to (14). Here, the object has a GN interpretation. On the other hand, the patient voice sentence in (15) is semantically identical to (16) and it continues to lack a GN interpretation.

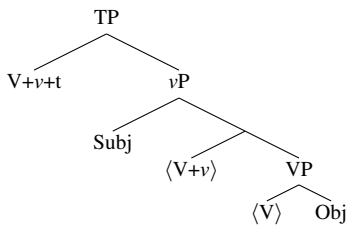
- (13) B⟨um⟩ili ang bata ng libro. (14) B⟨um⟩ili ng libro ang bata.
 AV.buy ANG child NG book AV.buy NG book ANG child
 ‘The child bought one or more book(s).’ ‘The child bought one or more book(s).’
- (15) B⟨in⟩ili ng bata ang libro. (16) B⟨in⟩ili ang libro ng bata.
 PV.buy NG child ANG book PV.buy ANG book NG child
 ‘The child bought the book.’ ‘The child bought the book.’

Therefore, Tagalog shares the semantic properties of traditional NI but not the syntactic properties. Due to this fact, I assert that Tagalog semantic incorporation (SI) where the term was first coined by Massam (2001).

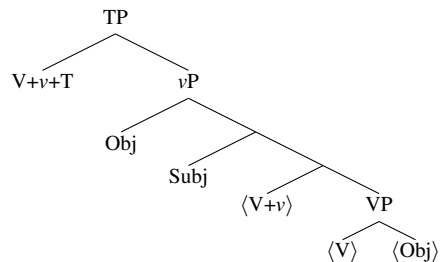
3. Structure and proposal

The structure of verb initiality can differ cross-linguistically, it can be the result of either head raising or phrasal raising. In Tagalog, both types of raising have been debated as well as subject lowering (Sabbagh 2014). Raising in some languages clearly demonstrates either head or phrasal raising but because Tagalog has free post-verbal word order, determining raising in Tagalog is a slightly more complex problem.

(17) Head raising, VSO



(18) Head raising, VOS



I hypothesize that there are two object copies, a lower copy and a higher copy and that the interpretation and pronunciation of either of these copies at LF and PF respectively has implications on voice and word order. For instance, interpretation of the higher object copy at LF, results in a specific object in a patient voice sentence. On the other hand, interpretation of the lower object copy at LF, results in an agent voice sentence. Pronunciation at PF determines word order. Pronouncing the lower object copy results in VSO word order and pronouncing the higher object copy results in VOS word order.

In summary, I hypothesize that voice is a morphological indicator of which copy of the object is interpreted at LF and word order is the result of which copy of the object is pronounced at PF.

1	LF	higher copy	patient VOS	2	LF	higher copy	patient VSO
	PF	higher copy			PF	lower copy	
3	LF	lower copy	agent VOS	4	LF	lower copy	agent VSO
	PF	higher copy			PF	lower copy	

Table 2: Possible outcomes due to interpretation of the object

The chart in table 2 concisely organizes the interaction of object copies with PF and LF. For example, in configuration 1, at LF the higher copy is interpreted and at PF the higher copy is also pronounced. This results in a patient voice sentence with VOS word order. The existence of a higher copy is assumed and in every configuration except 4, the higher copy is accessed at either LF, PF, or both. In configuration 4, the lower copy is used both at LF and PF. This raises doubt as to whether a higher copy exists at all. Edith Aldridge (p.c.) asserts that in these scenarios, the sentence undergoes VP raising to a [Spec TP] and that only the subject occupies [Spec *v*P]. The issue with this is that to say that sometimes objects move and sometimes they don't results in a much more complex system than to simply say that objects always move but are sometimes not accessed. For the present paper, I maintain that objects always move.

4. Review on incorporation

There exist various semantic analyses on NI, Chung and Ladusaw (2004), Farkas and de Swart (2003), Van Geenhoven (1998) and Dayal (1999) to name a few. I adopt the analysis by Chung and Ladusaw (2004). In this section, I outline the mechanics of this text and then I describe how it fits in with the Tagalog data.

4.1. Incorporation via modes of composition Chung and Ladusaw (2004)

The approach in this text is based on the distinction between two modes of composition: *Restrict* and *Specify*. Transitive predicates have two unsaturated argument positions and the way that these arguments are composed can determine whether there it is incorporated or not. Arguments that are composed with the predicate via *Restrict* are incorporated. Arguments that are composed with the predicate via *Specify* are not.

4.1.1. Saturation

Specify is a saturating mode of composition that is often considered the standard mode. *Function application* (FA) and *existential closure* (EC) are the two saturating operations that we will see in this paper. For instance, a transitive predicate such as *see* has the following denotation in (19) where the two variables, *y* and *x*, are placeholders for arguments. This verb is of semantic type $\langle e, \langle e, t \rangle \rangle$ and it has two

unsaturated positions.

$$(19) \quad \text{see} = \lambda y \lambda x [\text{see}'(y)(x)]$$

One way to saturate the predicate is via FA. In this operation, the predicate merges with an argument of semantic type e which yields a function of $\langle e, t \rangle$. At this point, the predicate has one degree of incompleteness left, see the second last line of (20). Next, the predicate merges with another argument of semantic type e which yields a truth value t , see the final line of (20).

$$(20) \quad \begin{aligned} & \text{Suzanne (s) sees Vesela (v).} \\ & = \text{FA}(\text{FA}(\lambda y \lambda x [\text{see}'(y)(x)], (v)), (s)) \\ & = \text{FA}(\lambda x [\text{see}'(v)(x)], (s)) \\ & = \text{see}'(v)(s) \end{aligned} \quad t$$

Saturation via EC is another method. Rather than satisfying a predicate's selectional requirements, EC removes the incompleteness of a predicate instead of an existential quantification over the argument (Chung and Ladusaw 2004:4). The application of EC to the same predicate (19) reduces it from type $\langle e, \langle e, t \rangle \rangle$ to $\langle e, t \rangle$. Additional application of EC reduces it to a truth value t , see (21).

$$(21) \quad \begin{aligned} & = \text{EC}(\text{EC}(\lambda y \lambda x [\text{see}'(y)(x)])) \\ & = \text{EC}(\lambda x (\exists y [\text{see}'(y)(x)])) \\ & = \exists x \exists y [\text{see}'(y)(x)] \end{aligned} \quad t$$

EC can be implemented freely or as a last resort but this distinction is language-specific (Chung and Ladusaw 2004:92). EC is freely available in Maori (Reinhart 1997, Winter 1997) but it is only available as a last resort in St'át'imcets (Kratzer 1998).

4.1.2. Restriction

Restrict is a non saturating mode of composition that restricts the contexts in which a statement can be true. Arguments that are property denoting, such as *dessert* in (22), are treated as a restrictive modifier of the predicate (Chung and Ladusaw 2004:6). For instance, the set of eating events is a superset of the set of dessert-eating event. The application of *Restrict* to this scenario is illustrated in (22).

$$(22) \quad \begin{aligned} & \text{Restrict}(\lambda y \lambda x [\text{eat}'(y)(x)], \text{dessert}') \\ & = \lambda x \lambda y [\text{eat}'(y)(x) \ \& \ \text{dessert}'(y)] \end{aligned}$$

A complete derivation is presented in (23) where the type of dessert is underspecified. In (23), after *Restrict* is applied, EC follows because no particular dessert is specified. Finally, the subject is computed via FA (Chung and Ladusaw 2004:7).

$$(23) \quad \begin{aligned} & \text{Paul (p) dessert-ate.} \\ & \text{FA}(\text{EC}(\text{Restrict}(\lambda y \lambda x [\text{eat}'(y)(x)], \text{dessert}')), (p)) \end{aligned}$$

$$\begin{aligned}
 &= \text{FA}(\text{EC}(\lambda y \lambda x [\text{eat}'(y)(x) \ \& \ \text{dessert}'(y)]), (p)) \\
 &= \text{FA}(\lambda x \exists y [\text{eat}'(y)(x) \ \& \ \text{dessert}'(y)], (p)) \\
 &= \exists y [\text{eat}'(y)(p) \ \& \ \text{dessert}'(y)]
 \end{aligned}$$

The authors stress that *Restrict* is non-saturating by contrasting (23) with (24). In (23), there is no specified dessert, however a derivation with *Restrict* is still possible even if the dessert is specified. If the type of dessert is indeed specified, following *Restrict*, the derivation need not undergo EC but FA twice as shown in (24). The outcome is a fully saturated predicate. The thing that Paul ate must have some dessert-like property, hence birthday cake is accepted but something like asparagus is not.

$$\begin{aligned}
 (24) \quad &\text{Paul } (p) \text{ dessert-ate birthday cake } (b). \\
 &\text{FA}((\text{FA}((\text{Restrict}(\lambda y \lambda x [\text{eat}'(y)(x)], \text{dessert}')), (b))), (p)) \\
 &= \text{FA}((\text{FA}((\lambda y \lambda x [\text{eat}'(y)(x) \ \& \ \text{dessert}'(y)]), (b))), (p)) \\
 &= \text{FA}((\lambda x [\text{eat}'(b)(x) \ \& \ \text{dessert}'(b)]), (p)) \\
 &= \text{eat}'(b)(p) \ \& \ \text{dessert}'(b)
 \end{aligned}$$

It is an important fact that semantic well-formedness does not always entail syntactic acceptability. That being said, while the above sentence is not a well-formed sentence of English, it is well-formed in a Chamorro sentence with similar compositional patterns.

4.1.3. Chamorro object incorporation

Chamorro, another Austronesian language, has object incorporation. What is interesting about Chamorro incorporation is that incorporating predicates can have both an incorporated object and another independent noun phrase which the authors call an *extra object* (Chung and Ladusaw 2004:104). *Birthday cake* in (24) is another example of an extra object. The following examples, (25) and (26), demonstrate object incorporation without and with an extra object, respectively. Both examples incorporate *ga'* 'pet' but the latter also has an object *kätu* 'cat'.

<p>(25) Män-gäi-[ga'] häm. agr.have-pet we 'We have pets.'</p>	<p>(26) Gäi-[ga'] yu' kätu, lao matai. agr.have-pet I cat but age.die 'I had a pet cat, but it died.'</p>
--	---

The authors make three points. First, the incorporated object is semantically incomplete and thus denotes a property rather than an individual. Therefore, they can only compose via *Restrict*. Second, the extra object is not a complement of V but an adjunct. Third, the incorporated object and the extra object are both composed with the internal argument of the verb. This is referred to as *multiple linking* (Chung and Ladusaw 2004:102).

Object incorporation in Chamorro is limited to verbs of possession, *gäi* 'have' (27) and *täi* 'not have' (28) (Chung and Ladusaw 2004:111). Despite this limitation, any common noun can be incorporated.

- (27) Gäi-[famagu'un] ädyu na palao'an.
agr.have-children that L woman
'That woman has children.'
- (28) Täi-[prublema] i próhimu.
agr.not.have-problem the fellow
'The guy had no problem.'

As an ergative language, transitive objects and intransitive subjects have matching morphology that is distinct from transitive subjects. Compare the agreement markers in (29) - (31) where the subjects are both 3rd person plural yet identified as *man* in (29) and *ma* in (30). Incorporating predicates are treated as intransitives. The agreement marker in (31) is the same as the one in (29).

- | | |
|--|--|
| <p>(29) Man-malagu.
agr-run
They ran.</p> | <p>(30) Ma-bäba i petta.
agr-open the door
'They opened the door.'</p> |
| <p>(31) Man-gäi-famagu'un.
agr-have-children
'They have children.'</p> | <p>(Chung and Ladusaw 2004:113)</p> |

The Chamorro extra object is particularly interesting because there is a mismatch between its structural and its semantic features. The extra object does not resemble a DP complement of V because it lacks some of the morphology and all of the syntax of one (Chung and Ladusaw 2004:125). On the other hand, the extra object is not a semantic adjunct but an argument. Extra objects are composed with the verb's internal argument and can saturate that argument (Chung and Ladusaw 2004:125).

As previously stated, incorporated objects are semantically incomplete. They are property denoting and are composed via *Restrict*. If there is an extra object, it can saturate the predicate via FA (Chung and Ladusaw 2004:132). When there is no extra object, the unsaturation of the internal argument is satisfied by EC. Both lead to a semantically complete predicate.

The saturation of predicates that have both an incorporated nominal and an extra object once again demonstrates multiple linking (Chung and Ladusaw 2004:145). In order for multiple linking to succeed, *Restrict* must be ordered first. If FA composes the extra object first, then the predicate becomes saturated and cannot be restricted.

- (32) Si Carmen gäi-[ga'] i ga'lagu.
Unm Carmen agr.have-pet the dog
'Carmen has the dog as pet.'
- (33) FA((FA((Restrict($\lambda y \lambda x$ [have'(y)(x)], pet')), dog')), Carmen')
= FA((FA(($\lambda y \lambda x$ [have'(y)(x) & pet'(y)]),dog')), Carmen')
= FA((λy [have'(d)(c) & pet'(d)]), Carmen')
= have'(d)(c) & pet'(d)

The analysis of Chamorro demonstrates the semantic incompleteness of incorporated objects such that extra objects may saturate the predicate afterwards. An appealing feature of Chung and Ladusaw (2004) is that *Restrict* can technically compose multiple arguments at the risk of overgenerating. This means that *Restrict* could potentially compose six arguments in language X. For instance, *Restrict* can compose two nominals with the verb as shown in the Chamorro example in (32). While this composition is acceptable in Chamorro, it is not acceptable in languages like Tagalog and English. Tagalog has incorporated nominals but it does not have extra objects. The authors suggest that it is not necessarily the responsibility of their approach but the responsibility of the syntax to determine whether an incorporating language may have an extra object or not (Chung and Ladusaw 2004:149).

4.2. Motivation for using Restriction and Saturation

Some other analyses of incorporation not discussed in the present paper look at Hindi (Dayal 1999, 2011), Hungarian (Farkas and de Swart 2003), and West Greenlandic (Van Geenhoven 1998). These various analyses of incorporation differ in part because the empirical phenomena differ cross-linguistically. Comparing incorporation in Tagalog with incorporation in these languages has facilitated selecting a framework for the analysis in section 5. For instance, Hungarian and Hindi can incorporate plural nominals where Tagalog cannot. In West Greenlandic, predicates are lexically specified to incorporate or not. In Tagalog, there is virtually no restriction specifying which verbs can incorporate and which cannot.

Despite some differences between Chamorro and Tagalog, Chung and Ladusaw (2004) is a good starting point for my analysis because it does not distinguish between singular or plural INs like in Hindi and Hungarian and it does not lead me to believe that incorporating predicates are any different from non-incorporating predicates like in West Greenlandic.

Chamorro is the only language of the group that can have both an IN and an extra object. Fortunately, the presence of an extra object in Chamorro is not problematic to my analysis because they are optional. An incorporating sentence without an extra object can still achieve completion as shown in (23).

Discourse transparency is another point of contention. Chamorro and West Greenlandic are discourse transparent such that the IN can serve as antecedent to a subsequent pronoun. Hungarian singular INs are discourse translucent such that they can antecede covert pronouns but not overt ones. Hungarian plural INs are discourse transparent. Hindi singular INs are discourse opaque and Hindi plural INs are discourse transparent. Tagalog INs are discourse transparent.

The approaches by Chung and Ladusaw (2004), Farkas and de Swart (2003), Van Geenhoven (1998), and Dayal (1999) have some principal similarities with respect to their treatment of incorporation. INs, for instance, are consistently described as property denoting and are treated as such in their semantic interpretations. With respect to number neutrality, languages that only incorporate morphologically singular nominals, such as West Greenlandic and Chamorro, have GN IN. When languages

have both morphologically singular and plural INs, like Hungarian and Hindi, singular INs are number neutral and plural INs are semantically plural.

Finally some languages have restrictions on the verbs and/or nouns that can be incorporated. Only existential verbs can incorporate in Chamorro but they are free to incorporate any nominal. Hindi verbs on the other hand are relatively free to incorporate but are restricted by the types of nominals that can be incorporated. West Greenlandic and Hungarian have even more freedom for incorporation. These restrictions need not be explained in the semantics but may simply be stipulated in the lexicon. Therefore, this leaves open the choice of using Restriction and Saturation for an analysis on incorporation in Tagalog which is presented in the next section.

5. Application of Restriction and Saturation to Tagalog

While there are multiple semantic frameworks from which I can apply an analysis of Tagalog SI, I choose to model mine after Chung and Ladusaw (2004). One important difference between Chamorro and Tagalog is that Chamorro incorporation displays the traditional characteristics in that it is manifested morpho-syntactically whereas Tagalog incorporation is only manifested semantically. Morphosyntactically, incorporation in these languages do not match but semantically, they are quite similar. Chung and Ladusaw (2004)'s account of incorporation is semantically based thus borrowing their analysis for my own on Tagalog is still well founded.

5.1. Setting up the problem

In this section, I set up my semantic analysis of Tagalog SI by fleshing out the characteristics of Tagalog. As previously stated, *ang* arguments do not have GN regardless if in subject position as in (1) or object position as in (4). On the other hand, some *ng*-arguments can have GN in special environments like the unmodified object in (3). GN does not arise, however, in modified objects as shown with the adjectival modifier in (8) and the plural particle in (9). *Ng* subjects never have GN.

Aldridge (2004, 2012) asserts that agent voice predicates are intransitive because they share the same *-um-* verbal morphology as other intransitive predicates, see (34). This assertion supports the argumentation that Tagalog is an ergative language where patient voice sentences are unmarked transitives and agent voice sentences are intransitive. This is an interesting observation that demonstrates the similarities of Tagalog incorporation with many other incorporating languages like Chamorro, Hungarian, and West Greenlandic.

- (34) D⟨um⟩ating ang babae.
intr.arrive ang woman
'The woman arrived.' (Aldridge 2012:192)

Finally, some languages like Chamorro and Hindi demonstrate a restriction on the class of verbs available to incorporate and/or the class of nouns available to be incorporated.

For Tagalog, theoretically, any transitive verb can combine with any non-pronominal argument to form an incorporating construction.

To summarize this section, INs in Tagalog are morphologically singular. Incorporating predicates have agent voice morphology which may indicate intransitivity following Aldridge (2004, 2012). Finally any non-pronominal unmodified argument can incorporate with a verb in agent voice.

5.2. Semantic lexicon

Before delving into the application of Chung and Ladusaw (2004)'s framework to Tagalog, it is imperative to define the lexicon. Of particular importance are the semantic denotations of *ang* and *ng* which, as it has been already established, have various semantic contrasts. In section 2.2 and 2.3 I reviewed Paul et al. (In Press) which demonstrates that, apart from modification by *isang*, there is virtually no distinction between *ang* and *ng*. Therefore, any interpretative differences are driven by structure and not by these particles themselves. However, because the facts about voice, subjects and objects, and modification are quite complicated, I initially simplify these facts and treat *ang* as a definite determiner while analyzing *ng*. (See (35) for the denotation.)

$$(35) \quad \llbracket \textit{ang} \rrbracket = \lambda P_{\langle e,t \rangle} \iota x. P(x) \quad \langle \langle e, t \rangle, e \rangle$$

$$(36) \quad \llbracket \textit{ng} \rrbracket = \emptyset$$

GN nominals are unmodified *ng* objects which are interpreted within VP at LF and require EC to bind them. Conversely, *ang* objects are interpreted outside of VP and because of this are not bound by EC.

Subjects are generated outside of VP, therefore, they are never interpreted within VP which means that they are not traditionally bound by EC. *Ng* subjects denote individuals but licensing them is problematic because *ng* has no licensing power. One way to remedy this is to treat *ng* subjects similar to *ng* objects such that they require some closure operation like EC but without the incorporation interpretation that is typically associated by EC. In the derivations, I denote this EC-like operation as XC.

In addition to the denotations of *ang* and *ng*, I provide the following additional denotations in (37) which will be used in the following derivations.

$$(37) \quad \begin{aligned} \llbracket (\textit{some transitive verb}) \rrbracket &= \lambda y_e \lambda x_e. [(\textit{some transitive verb})'(y)(x)] \langle e, \langle e, t \rangle \rangle \\ \llbracket (\textit{some noun}) \rrbracket &= \lambda x_e. \textit{some noun}(x) \quad \langle e, t \rangle \end{aligned}$$

5.3. Agent voice object = Incorporation

Despite differences in composition, both sentences with incorporation and sentences without it can achieve completeness. In a sentence such as (13), the unmodified *ng* object is incomplete and must enter the derivation via *Restrict*. This can be seen as some form of modifying the event such that the only possible things that were bought

were things with book-like properties. This operation does not saturate the verb and leaves it with two argument slots still open. Because Tagalog is not like Chamorro, it does not have both an IN and an extra object. The IN is bound by EC at this stage, thus blocking the possibility for an extra object. This leaves one argument slot available for interpretation. The subject, which is *ang* marked, enters the derivation via FA thereby completing it.

$$\begin{aligned}
 (38) \quad &= \text{FA}(\text{EC}(\text{Restrict}(\lambda y \lambda x [\text{buy}'(y)(x)], (\lambda z. \text{book}(z))), (\iota v. \text{child}(v)))) \\
 &= \text{FA}(\text{EC}(\lambda y \lambda x [\text{buy}'(y)(x) \ \& \ \text{book}'(y)], (\iota v. \text{child}(v)))) \\
 &= \text{FA}(\lambda x \exists y [\text{buy}'(y)(x) \ \& \ \text{book}'(y)], (\iota v. \text{child}(v))) \\
 &= \exists y [\text{buy}'(y)(\iota v. \text{child}(v)) \ \& \ \text{book}'(y)] \quad t
 \end{aligned}$$

5.4. Patient voice object = non-incorporation

Objects of sentences without incorporation are entity denoting. In (15), the object which is *ang* marked saturates one argument position of the verb via FA. The semantically incomplete subject enters the derivation via *Restrict* and then requires XC to bind it.

$$\begin{aligned}
 (39) \quad &= \text{EC}(\text{Restrict}((\text{FA}(\lambda y \lambda x [\text{buy}'(y)(x)], (\iota z. \text{book}(z))), (\lambda v. \text{child}(v)))) \\
 &= \text{EC}(\text{Restrict}(\lambda x [\text{buy}'(\iota z. \text{book}(z))(x)], (\lambda v. \text{child}(v)))) \\
 &= \text{EC}(\lambda x [\text{buy}'(\iota z. \text{book}(z))(x) \ \& \ \text{child}(x)]) \\
 &= \exists x [\text{buy}'(\iota z. \text{book}(z))(x) \ \& \ \text{child}(x)] \quad t
 \end{aligned}$$

5.5. The effect of modification

Another context in which a sentence does not have incorporation is one in which a *ng* object is modified. Modified *ng* objects do not have GN and are, therefore, not incorporated. Because *ng* has no semantic value, these arguments must get their entity denoting quality elsewhere. This comes from the modifier whether it is *isang* ‘one’, *pulang* ‘red’, *mga* ‘(plural)’, or something else.

Dayal (2004) and Mathieu (2012) describe a phenomenon in Hindi and French respectively which is called *licensing by modification*. Licensing by modification is described by Dayal (2004:217) as: “[a phenomenon where] an unacceptable noun phrase is redeemed by the presence of a modifier or an unavailable reading for a noun phrase is made available with the addition of a modifier”. The modifier changes a nominal of semantic type $\langle e, t \rangle$ to one of type $\langle \langle e, t \rangle, t \rangle$. The modifier raises to the determiner domain where, with *de*, it can be distributed between a head and a specifier (Mathieu 2012:2).

In French, there are environments where unmodified noun phrases are not well-formed but are acceptable when modified, see example (40). Mathieu (2012) explains that the modifier raises into the domain of the determiner via phrasal movement. Articles are comprised of a head and a specifier in modified contexts Damourette and Pichon (1911-1940). In such contexts, *de* occupies a head position while an

adjective or *les* occupies the specifier position. In Tagalog, the specifier position can be occupied by an adjective or some number specifier such as *isang* or the plural particle, *mga*.

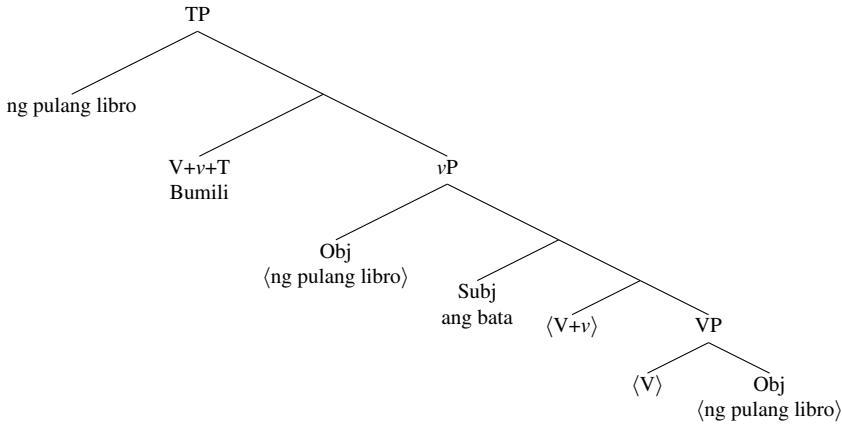
- (40) J'ai lu de *(bons) romans l'été dernier.
 1sg.have read *de* (good) novels the.summer last
 'I read some novels last summer.' (Mathieu 2012:1)

In Tagalog, the modifier has a functional head that triggers licensing. This functional head has various overt forms as well as a covert form. Modifiers that end in a vowel such as *pula* 'red' and *isa* 'one' have a suffixal linker *-ng* and modifiers with codas such as *mahal* 'expensive' have a free particle *na* in between the modifier and the nominal, see (41). The plural particle, *mga*, has a covert functional head. Robert Truswell (p.c.) suggests the denotation in (42) for this functional head. Further, the derivation in (43) demonstrates the application of this functional head with a nominal. The functional head, *-ng* in this example, takes the object as its argument.

- (41) mahal na libro
 expensive NA book
 'expensive book'
- (42) $\llbracket -ng \rrbracket = \lambda P \lambda M \lambda Q \exists x. [M(P)(x) \ \& \ Q(x)]$
 where $P = \langle e, t \rangle$ (nominal)
 $M = \langle \langle e, t \rangle, \langle e, t \rangle \rangle$ (modifier or numeral)
 $Q = \langle e, t \rangle$ (verbal property that the NP is a property of)
- (43) $\llbracket -ng \ libro \rrbracket = \llbracket -ng \rrbracket (\llbracket libro \rrbracket)$
 $= \lambda P \lambda M \lambda Q \exists x. [M(P)(x) \ \& \ Q(x)] (\lambda x. book(x))$
 $= \lambda M \lambda Q \exists x. [M(\lambda x. book(x))(x) \ \& \ Q(x)]$ (p.c. Robert Truswell)

In the next step, the modifier acts as a generalized quantifier when merged with the functional head and the nominal. This modified *ng* object is of semantic type $\langle \langle e, t \rangle, t \rangle$ and it raises to a position above the subject, see (45). The verb merges with the subject and this yields an argument of semantic type $\langle e, t \rangle$, see (46). The raised modified *ng* object serves as a function that takes the merged verb and subject as its argument, see (44).

(44)



(45)

$$\begin{aligned}
 & \llbracket \text{pulang libro} \rrbracket && \text{(p.c. Robert Truswell)} \\
 & = \llbracket -ng \text{ libro} \rrbracket (\llbracket \text{pula} \rrbracket) \\
 & = (\lambda M \lambda Q \exists x. [M(\lambda x. \text{book}(x))(x) \ \& \ Q(x)])(\lambda P \lambda x. \text{red}(x) \ \& \ P(x)) \\
 & = \lambda Q \exists x. [\lambda P \lambda x. \text{red}(x) \ \& \ P(x) (\lambda x. \text{book}(x))(x)] \ \& \ Q(x) \\
 & = \lambda Q \exists x. [(\lambda x. [\text{red}(x) \ \& \ \text{book}(x)](x))(x) \ \& \ Q(x)] \\
 & = \lambda Q \exists x. [\lambda x. [\text{red}(x) \ \& \ \text{book}(x)](x) \ \& \ Q(x)] \\
 & = \lambda Q \exists x. [\text{red}(x) \ \& \ \text{book}(x) \ \& \ Q(x)] && \langle \langle e, t \rangle, t \rangle
 \end{aligned}$$

(46)

$$\begin{aligned}
 & \llbracket \text{bumili ang bata} \rrbracket \\
 & = \llbracket \text{bumili} \rrbracket (\llbracket \text{ang bata} \rrbracket) \\
 & = \lambda x. \text{bought}(\text{the child})(x) && \langle e, t \rangle
 \end{aligned}$$

(47)

$$\begin{aligned}
 & \llbracket \text{bumili ang bata ng pulang libro} \rrbracket \\
 & = \llbracket \text{ng pulang libro} \rrbracket (\llbracket \text{bumili ang bata} \rrbracket) \\
 & = (\lambda Q \exists x. [\text{red}(x) \ \& \ \text{book}(x) \ \& \ Q(x)])(\lambda x. \text{bought}(\text{the child})(x)) \\
 & = \exists x. [\text{red}(x) \ \& \ \text{book}(x) \ \& \ \text{bought}(\text{the child})(x)] && t
 \end{aligned}$$

5.6. Returning to ANG

One issue with my analysis is the application of EC over both *ng* objects as well as *ng* subjects, which I refer to earlier as XC. Traditionally, EC only binds nominals within VP which is problematic because *ng* subjects are actually VP-external. While my analysis breaks with the traditional treatment of EC, what it actually needs is some closure operation that can bind *ng* subjects in a similar fashion as EC without some incorporated meaning.

As shown in Paul et al. (In Press), unmodified *ang* looks specific however *ang isang* looks nonspecific. *ng* subjects also appear specific while, *ng isang* looks nonspecific. This removes any contrast between *ang* and *ng*. My analysis treats *ng* as semantically vacuous but I also believe that *ang* is semantically vacuous.

I assert that *ang* forces VP-external interpretations whereas *ng* forces interpretation before movement. This means that *ng* objects are interpreted VP-internally and *ng* subjects are interpreted VP-externally. The interpretative differences that arise are the result of different closure operations that apply VP-internally and VP-externally.

6. Conclusion

For the present paper, I have identified some limitations on GN that were not previously described, primarily that GN nominals must be unmodified *ng* objects. This has led to my argumentation that Tagalog has incorporation. I have described Tagalog as a language that shares the semantic properties of noun incorporation but not the structural properties. Incorporation in Tagalog depends on two things: first, the predicate must be in agent voice, and second, the object must be unmodified. Achieving these two things yields a GN interpretation on the object. Traditional incorporation requires the incorporated nominal to be morphologically attached or at least close to the verb. Tagalog has free post-verbal word order and regardless of whether a sentence has VOS or VSO word order the semantics is unchanged. Therefore, Tagalog has semantic incorporation.

My analysis leans on the analysis by Chung and Ladusaw (2004) where INs are composed with the verb via *Restrict*. This contrasts with unincorporated nominals that are composed with the verb via FA. The semantic application of the Tagalog data to Chung and Ladusaw (2004) demonstrates the well-formedness behind labelling Tagalog as an incorporating language. The theoretical implications of this work are that it contributes to the expanding literature on incorporation and it introduces a new flavour of incorporation, voice dependency.

References

- Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.
- Aldridge, Edith. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122:192–203.
- Allen, Barbara, Donna Gardiner, and Donald Frantz. 1984. Noun incorporation in Southern Tiwa. *International Journal of American Linguistics* 50:292–311.
- Baker, Mark C. 1988. *Incorporation*. Chicago: University of Chicago Press.
- Carlson, Greg. 2006. The meaningful bounds of incorporation. In *Non-definiteness and plurality*. John Benjamins Publishing Company.
- Chung, Sandra, and William A Ladusaw. 2004. *Restriction and saturation*. MIT press.
- Corbett, Greville G. 2000. *Number*. Cambridge University Press.
- Damourette, Jacques, and Édouard Pichon. 1911-1940. *Des mots à la pensée: essai de Grammaire de la langue française*, volume 7. Paris: Dartrey.
- Dayal, Veneeta. 1999. Bare NP's, Reference to Kinds, and Incorporation. In *Proceedings of SALT*, volume 9, 34–51.
- Dayal, Veneeta. 2004. Licensing by modification. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies* 217–238.
- Dayal, Veneeta. 2011. Hindi pseudo-incorporation. *Natural Language & Linguistic Theory* 29:123–167.
- Farkas, Donka F., and Henriëtte de Swart. 2003. *The semantics of incorporation: From argument structure to discourse transparency*. CSLI Publications.

- Johns, Alana. 2007. Restricting noun incorporation: Root movement. *Natural Language & Linguistic Theory* 25:535–576.
- Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide-scope indefinites? In *Events and grammar*, ed. Susan Rothstein, 163–198. Dordrecht: Kluwer.
- Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language & Linguistic Theory* 19:153–197.
- Massam, Diane. 2009. Noun incorporation: Essentials and extensions. *Language and Linguistics Compass* 3:1076–1096.
- Mathieu, Éric. 2012. Licensing by modification: The case of french de nominals. *Journal of Linguistics* 48:389–426.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 847–894.
- Paul, Ileana, Key Cortes, and Lareina Milambiling. In Press. Definiteness without D: The case of *ang* and *ng* in Tagalog. MS., to appear in the Canadian Journal of Linguistics.
- Postal, Paul. 1962. Some syntactic rules of Mohawk. Doctoral Dissertation, Yale University.
- Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335–397.
- Sabbagh, Joseph. 2014. Word order and prosodic-structure constraints in tagalog. *Syntax* 17:40–89.
- Schachter, Paul, and Fe T Otanes. 1972. *Tagalog reference grammar*. University of California Press.
- Seiter, W. 1980. *Studies in Niuean Syntax*. Garland, New York.
- Van Geenhoven, Veerle. 1998. *Semantic incorporation and indefinite descriptions: Semantic and syntactic aspects of noun incorporation in west greenlandic*. CSLI publications.
- Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20:399–467.

AN AGENT ADVANTAGE
IN TAGALOG RELATIVE CLAUSE COMPREHENSION*

Nozomi Tanaka
University of Hawai'i at Mānoa
nozomit@hawaii.edu

William O'Grady
University of Hawai'i at Mānoa
ogradey@hawaii.edu

Kamil Deen
University of Hawai'i at Mānoa
kamil@hawaii.edu

Chae-Eun Kim
Korea University
alohachaeun@gmail.com

Ryoko Hattori
University of Hawai'i at Mānoa
rhattori@hawaii.edu

Ivan Paul M. Bondoc
University of the Philippines, Manila
imbondoc@up.edu.ph

Jennifer U. Soriano
University of the Philippines, Manila
jusoriano@up.edu.ph

Previous studies of Tagalog report that children were more successful at producing agent relative clauses than theme relative clauses. The current study documents a similar asymmetry in Tagalog-speaking children's comprehension of relative clauses.

1. Introduction

A signature feature of Tagalog grammar is its Philippine-type focus system, in which a verbal affix signals the thematic role of the privileged syntactic argument (PSA), a nominal usually marked by the case prefix (focus marker) *ang*. In (1), the *-um-* infix in the verb indicates that the NP bearing the marker *ang* is the agent. In (2), on the other hand, the infix *-in-* indicates that the PSA is the patient/theme.

* We thank the University of the Philippines (UP) Child Development Center and the UP Department of Linguistics for allowing us to conduct the study. We are also grateful to three anonymous reviewers and the participants at AFLA 22 for valuable comments. This study was supported by an Arts and Sciences Student Research Award at the University of Hawai'i given to the first author in 2014.

- (1) AGENT FOCUS
 H<um>a~hábol aŋ láláke naŋ babáe.
 <AF>IPFV~chase FOC man NFOC woman¹
 ‘The man is chasing a/the woman.’
- (2) THEME FOCUS
 H<in>a~hábol naŋ láláke aŋ babáe
 <TF>IPFV~chase NFOC man FOC woman
 ‘A/The man is chasing the woman.’

These contrasts are also relevant to the syntax of relative clauses. As illustrated in (3) and (4), a PSA is relativizable.

- (3) AGENT RELATIVE CLAUSE WITH AGENT FOCUS
 láláke=ŋ [h<um>a~hábol naŋ babáe]
 man=L <AF>IPFV~chase NFOC woman
 ‘(the) man who is chasing a/the woman’
- (4) THEME RELATIVE CLAUSE WITH THEME FOCUS
 babáe=ŋ [h<in>a~hábol naŋ láláke]
 woman=L <TF>IPFV~chase NFOC man
 ‘(the) woman who a/the man is chasing’

In contrast, other arguments cannot be relativized, as shown in (5) and (6).

- (5) AGENT RELATIVE CLAUSE WITH THEME FOCUS
 ?? láláke=ŋ [h<in>a~hábol aŋ babáe]
 man=L <TF>IPFV~chase FOC woman
 ‘(the) man who is chasing the woman’
- (6) THEME RELATIVE CLAUSE WITH AGENT FOCUS
 * babáe=ŋ [h<um>a~hábol aŋ láláke]
 woman=L <AF>IPFV~chase FOC man
 ‘(the) woman who the man is chasing’

While previous studies treat the pattern in (5) as ungrammatical (e.g. Aldridge 2004), it is apparently acceptable for many speakers, even though (3) is the preferred way to relativize an agent argument. It is not yet clear whether this is a dialectal difference or a generational difference.

Cross-linguistic research has shown that subject (agent) relative clauses are easier to produce, comprehend, and acquire than direct object (theme) relative

¹ List of abbreviations: AF = agent focus; FOC = focus marker; IPFV = imperfective; L = linker; NFOC = non-focus marker; TF = theme focus

clauses in many languages,² including Dutch (e.g. Frazier 1987), English (e.g. King and Just 1991; Diessel and Tomasello 2005), German (e.g. Schriefers, Friederici, and Kuhn 1995; Diessel and Tomasello 2005), French (Holmes and O'Regan 1981), Greek (Stavrakaki 2001), Hebrew (Friedmann, Belletti, and Rizzi 2009), Jakarta Indonesian (Tjung 2006), Japanese (e.g. Kawashima 1980; Ishizuka 2005; Miyamoto and Nakamura 2003), Korean (e.g. Cho 1999), Mandarin Chinese (e.g. Hsu, Hermon, and Zukowski 2009; Vasishth et al. 2013), Persian (Rahmany, Marefat, and Kidd 2011), and Swedish (Håkansson and Hansson 2000). But what about Tagalog?

Based on data from a picture-based elicited production task, Tanaka and colleagues (to appear) report that, although children and adults prefer theme focus patterns in basic clauses, children perform better on agent focus patterns when tested on relative clauses. However, we know from other studies, such as Gutierrez-Mangado & Ezeizabarrena (2012), that comprehension and production may yield different results. We therefore set out to investigate whether an agent advantage for Tagalog relative clauses also shows up in comprehension.

2. Method

2.1. Participants

Participants were 15 children (7 female; 8 male) from the age of 3;0 to 5;4 (mean age 4;3). We further divided the children into three age-related groups: three-year-olds (ages of 3;0–3;4, mean age 3;2, $n = 3$), four-year-olds (ages of 3;9–4;5, mean age 4;2, $n = 7$), and five-year-olds (ages of 4;8–5;4, mean age 5;0, $n = 5$). 15 adult speakers of Tagalog were also tested as control.

2.2. Materials

Test items contained a panel of two pictures, each illustrating a different version of the same action (e.g., a girl chasing a boy or a boy chasing a girl, as in Figure 1). The pictures were accompanied by an auditory description given in the form of either an agent relative clause or a theme relative clause.

² We take no position here on the grammatical status of agent and theme arguments in the various focus patterns of Tagalog—a matter on which there is some disagreement in the literature. Fortunately, this issue is independent of the question of whether which relative clause pattern is easier to comprehend and/or produce.

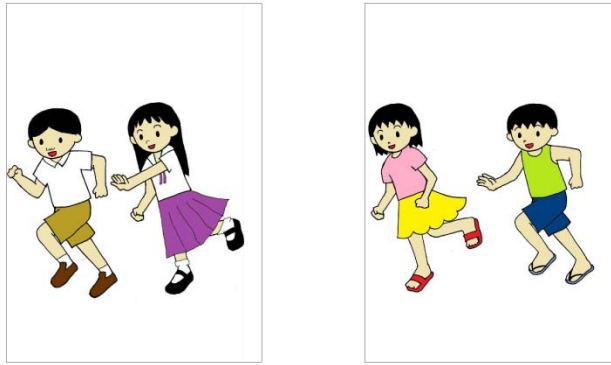


Figure 1. Sample item involving “chasing”.

If the participants heard the agent relative clause in (7) while viewing Figure 1, for example, the correct response was to select the girl in the picture on the left. In contrast, if they heard the theme relative clause in (8), they should select the girl in the picture on the right.

- (7) $laláke=\eta$ $[h<um>a\sim hábol$ $na\eta$ $babáe]$
 $man=L$ $<AF>IPFV\sim chase$ $NFOC$ $woman$
 ‘(the) man who is chasing a/the woman’
- (8) $babáe=\eta$ $[h<in>a\sim hábol$ $na\eta$ $laláke]$
 $woman=L$ $<TF>IPFV\sim chase$ $NFOC$ man
 ‘(the) woman who a/the man is chasing’

The same five verbs were used for both relative clause types (*basá?* ‘wet’, *búhat* ‘carry’, *hábol* ‘chase’, *túlak* ‘push’, *yákap* ‘hug’), giving us ten items in total. Because the verb is the disambiguating region for each relative clause (thanks to its focus affix), its occurrence was used as the starting point for measuring reaction time. In order to ensure comparability across test items, the stimuli were manipulated so that the onset of the verb always appeared at the 100 ms mark.

2.3. Procedure

With the use of four practice items, participants were first trained to select the character of their choice, and not the whole picture. The experiment with adults was carried out with the help of MouseTracker (Freeman and Ambady 2010), which recorded the movements of the computer mouse as participants used it to click on the character of their choice. The experiment with children involved a paper-and-pencil task, in which they were instructed to point at the character of their choice and then circle it with a pencil. We measured accuracy (i.e., the selection of the correct picture) for adults and children, as well as reaction times (i.e., how long it took to click on a character) for adults.

2.4. Analysis

We classified participants' responses with the help of the answer types illustrated in Figures 2 and 3. For the agent relative clause 'the girl who is chasing the boy', for instance, the girl on the left in Figure 2 is coded as the target referent—the agent. Choice of the other girl (the theme) was categorized as a reversal error, as the thematic role of the girl is reversed (from agent to theme). Responses that involved pointing to one of the boys were coded as head errors since the head modified by the relative clause must be a girl.

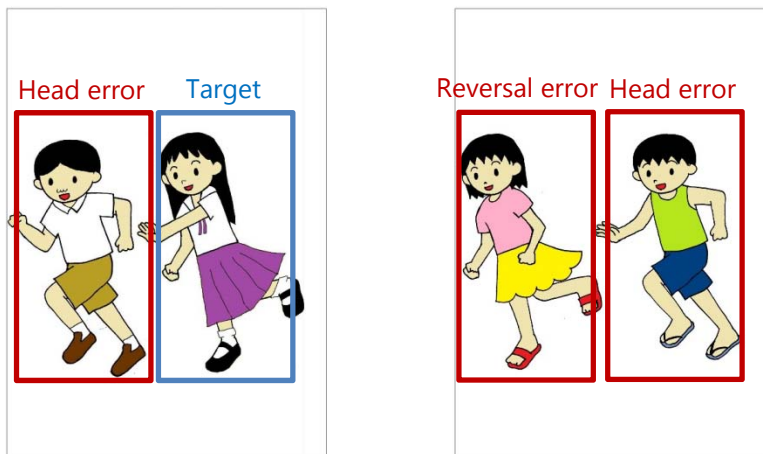


Figure 2. Response types for an agent RC 'the girl who is chasing the boy'.

For theme relative clause items, the girl on the right in Figure 3 was coded as the target referent—the theme argument. Selection of the other girl, who is the agent of the chasing action, was categorized as a reversal error. Selection of a boy was coded as a head error.

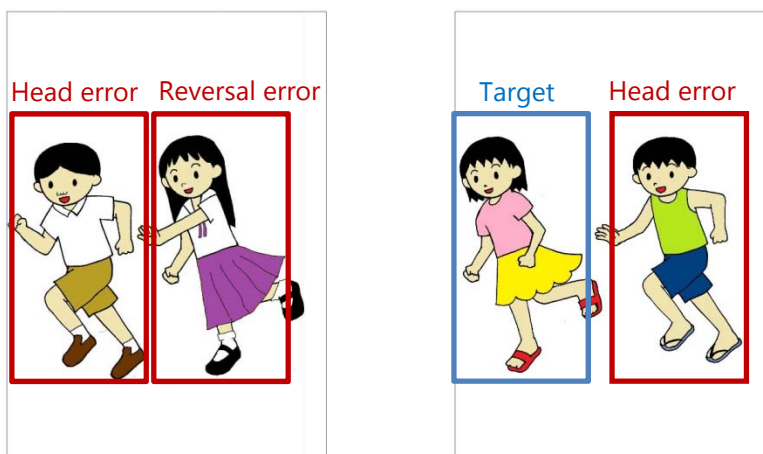


Figure 3. Response types for a theme RC 'the girl who the boy is chasing'.

3. Results

3.1. Adults

Adults clicked on the target referent 96.47% of the time when the prompt was an agent relative clause, and 97.65% of the time when the prompt was a theme relative clause. There was no significant difference between the two relative clause types in this regard.

Reaction times were calculated based on the target responses and were measured in two different ways: how long it took for the participants to initiate mouse movement (initial times³) and how long it took them to click (end times). Because the place in the picture where it was appropriate to click the mouse differed across the items, we also calculated normalized reaction times to average the distance between the initial location of the mouse and the location of the click.⁴ The mean values of these three measurements are shown in Table 1. No significant difference was found in any of these three measurements for the two types of relative clauses.

	Agent relative clauses	Theme relative clauses
Initial time	742	702
End time	3801	3898
Normalized reaction time	3142	3479

Table 1. Mean reaction times of adults (ms).

3.2. Children

Although three-year-olds and four-year-olds were generally poor at understanding either type of relative clauses, the data from five-year-olds showed a large advantage for agent relative clauses (72.4% correct) over theme relative clauses (28.6% correct). Figure 4 shows children's accuracy rates by age group.

³ The initial times prior to the verb onset at 100 ms were excluded as mouse movements during this period were not initiated with the intention of clicking.

⁴ Normalized reaction times were calculated as the duration between the initial time and the end time divided by the distance between the initial location of the mouse and the location of the click.

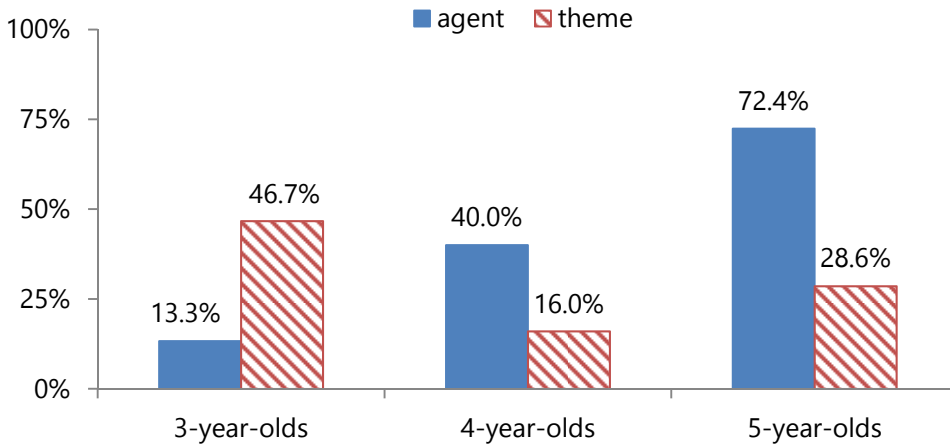


Figure 4. Accuracy rates of the children.

We also conducted an error analysis on the 5-year-olds' responses based on the error categories outlined in section §2.4.

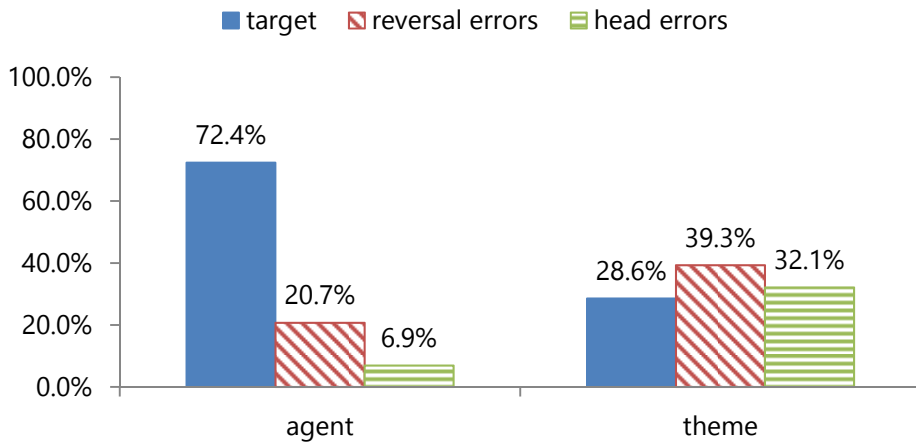


Figure 5. Response types of 5-year-olds.

The most frequent non-target responses involved reversals in which the participants switched the thematic role of the relativized element. Crucially, this error occurred almost twice as often in theme relative clauses as in their agent counterparts (39.3% versus 20.7%). That is, more theme relative clauses were interpreted as if they were agent relative clauses than vice versa.

4. Discussion

In sum, our study found no asymmetry between agent relative clauses and theme relative clauses in adult comprehension; however, children showed better accuracy

in the comprehension of agent relative clauses. Moreover, their most common errors were reversals in which they switched the thematic role of the relativized element. This happened more frequently with theme relative clauses, which means that more theme relative clauses were turned into agent relative clauses than vice versa, pointing to an agent relative clause advantage.

It is not clear yet whether the results from adults truly show a genuine symmetry in the production and comprehension of Tagalog relative clauses, or whether the tasks were not sensitive enough to uncover a difference. For this reason, we are planning a comparative study of English and Japanese using the same materials to see whether our tasks uncover an agent advantage in languages that are known to have such an asymmetry. In addition, we are planning to use other experimental techniques, such as self-paced reading and eye-tracking, to measure the production and comprehension of Tagalog relative clauses by adults.

4.1. Another interpretation

Although we have tentatively concluded that children's comprehension of Tagalog relative clauses manifest an agent advantage, there may be another way to explain our results.⁵ As noted in §1, many speakers apparently accept the relative clause pattern in (5), repeated as (9) below, despite claims to the contrary in previous research. That is, it is possible to relativize a non-PSA agent in a theme focus clause (Ceña and Nolasco 2011; Ceña and Nolasco 2012).

- (9) ?? *laláke=ŋ* [h<in>a~hábol aŋ babáe]
man=L <TF>IPFV~chase FOC woman
'(the) man who is chasing the girl'

Crucially, though, it is not possible to relativize a non-PSA theme in an agent focus clause. The following example repeats (6).

- (10) * *babáe=ŋ* [h<um>a~hábol aŋ laláke]
woman=L <AF>IPFV~chase FOC man
'(the) woman who the man is chasing'

This means that whereas the agent focus affix is an early and unambiguous cue of the thematic role of the relativized element (the 'head') of a relative clause, this is not true for the theme focus affix: relative clauses of this type could have a relativized agent or a relativized theme. In order to interpret these patterns, participants would have to wait until they encounter the other argument in the sentence in order to determine (from its case) whether it is the agent or the theme

⁵ We thank Henrison Hsieh, Aldrin Lee, and Chris Sundita for bringing this to our attention.

- Diessel, Holger, and Michael Tomasello. 2005. A new look at the acquisition of relative clauses. *Language* 81:882–906.
- Frazier, Lyn. 1987. Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory* 5:519–559.
- Freeman, Jonathan B., and Nalini Ambady. 2010. MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods* 42:226–241.
- Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119:67–88.
- Gutierrez-Mangado, M. Juncal, and Maria-josé Ezeizabarrena. 2012. Asymmetry in child comprehension and production of Basque subject and object relative clauses. *BUCLD 36 Proceedings Supplement*, eds. Alia Biller, Esther Chung, and Amelia Kimball.
- Håkansson, Gisela, and Kristina Hansson. 2000. Comprehension and production of relative clauses: A comparison between Swedish impaired and unimpaired children. *Journal of Child Language* 27:313–33.
- Holmes, V. M., and J.K. O'Regan. 1981. Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior* 20:417–430.
- Hsu, Chun-Chieh Natalie, Gabriella Hermon, and Andrea Zukowski. 2009. Young children's production of head-final relative clauses: Elicited production data from Chinese children. *Journal of East Asian Linguistics* 18:323–360.
- Ishizuka, Tomoko. 2005. Processing relative clauses in Japanese. *UCLA Working Papers in Psycholinguistics* 2:135–157.
- Kawashima, Michiko. 1980. The acquisition of Japanese relative clauses. Master's thesis, University of Alberta.
- King, Jonathan, and Marcel Adam Just. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30:580–602.
- Miyamoto, Edson T., and Michiko Nakamura. 2003. Subject/object asymmetries in the processing of relative clauses in Japanese. *WCCFL 22: Proceedings of the 22nd West Coast Conference on Formal Linguistic*, eds. Gina Garding, and Mimu Tsujimura, 342–355. Somerville, Ma: Cascadilla Press.
- Rahmany, Ramin, Hamideh Marefat, and Evan Kidd. 2011. Persian speaking children's acquisition of relative clauses. *European Journal of Developmental Psychology* 8:367–388.
- Schriefers, Herbert, Angela D. Friederici, and Katja Kühn. 1995. The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language* 34:499–520.
- Stavroulaki, Stavroula. 2001. Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language* 77:419–31.

The Proceedings of AFLA 22

- Tanaka, Nozomi, William O'Grady, Kamil Deen, Chae-Eun Kim, Ryoko Hattori, Ivan Paul M. Bondoc, and Jennifer U. Soriano. To appear. Split focus preferences in Tagalog: Evidence from child language. *AFLA XXI: Proceedings of the 21st Annual Meeting of Austronesian Formal Linguistics Association*, eds. Amber Camp, Claire Stabile, Nozomi Tanaka, and Yuko Otsuka. Canberra: Asia-Pacific Linguistics.
- Tjung, Yassir. 2006. The formation of relative clauses in Jakarta Indonesian: A subject-object asymmetry. Doctoral dissertation, University of Delaware.
- Vasishth, Shravan, Zhong Chen, Qiang Li, and Gueilan Guo. 2013. Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS ONE* 8.10:1-15.

DIAGNOSING VP-ELLIPSIS IN JAVANESE: EVIDENCE FOR A NON-MOVEMENT AND A MOVEMENT ACCOUNT*

Jozina Vander Klok

University of British Columbia, Max Planck Institute for Evolutionary Anthropology

jozina.vanderklok@mail.mcgill.ca

VP-ellipsis, a rare phenomenon cross-linguistically, has not been studied in Javanese. The main issue of this paper is to therefore investigate putative VP-ellipsis. Based on Goldberg's (2005) diagnostics, I show that alternative strategies such as stripping are not tenable for VP-ellipsis in Javanese and that it shares similar properties to VP-ellipsis in English. I conclude that VP-ellipsis is attested in this Western Malayo-Polynesian language. Further, I propose that VP-ellipsis in Javanese must involve two different strategies dependent on the environment. I suggest that VP-ellipsis in islands, coordinated CPs, and non-coordinated CPs (same speaker) involve a non-movement account, and VP-ellipsis in non-coordinated CPs (different speaker) and embedded clauses involve a movement account. Evidence for this partition is based on parallel behavior with VP-topicalization, where VP-topicalization feeds VP-ellipsis in the movement account.

1. Introduction

The central issue of this paper is to investigate putative VP-Ellipsis (VPE) in Javanese, as in examples like (1).¹ These types of propositions in Javanese appear to have hallmark features of VPE. That is, a licensing condition must be met; the auxiliary *ape* 'FUT' must be present in (1). The ellipsis site must also meet an identity condition. That is, the ellipsis site can only be understood as the same as its antecedent; in this case, as *melbu goa Maharani* 'enter the Maharani cave'.²

* *Matur suwun* to my consultants in Paciran (Mbak Titis, Mbak Fina, Mbak Ulum, and Mas Sari) and my consultants in Semarang (Mbak Wuri and Mbak Anna) for sharing your language and culture with me. Thank you also to the AFLA 22 participants for feedback, especially Ileana Paul, Norvin Richards, and Lisa Travis. This project is partly funded by SSHRC #756-2012-0648. All errors are my own.

¹ I use 'VPE' to refer to ellipsis of a verbal category, setting aside issues of its specific category.

² I follow the Leipzig glossing rules; additional glosses are as follows: AV 'actor voice', CIRC 'circumstantial modality', DEON 'deontic modality', EXP.PRF 'experiential perfect', NEC 'necessity', and POS 'possibility'.

- (1) Bu Haris gak ape melbu goa Maharani,
Mrs. Haris NEG FUT enter cave Maharani
tapi bojo-ne *(ape) [~~melbu goa Maharani~~].
but spouse-POSS FUT
'Bu Haris won't go into the Maharani cave, but her husband will.'

Despite these traits, VP-ellipsis can easily be misanalyzed as a different type of ellipsis, such as stripping (TP-ellipsis) or the combination of gapping (verb-ellipsis) and object pro-drop. Further, VP-ellipsis itself is a rare cross-linguistic phenomenon (Goldberg 2005).³ For these reasons, VPE should be identified using an explicit set of cross-linguistic diagnostics (Goldberg 2005:7). Using these diagnostics against a wider language database will then allow for refinement of them.

This paper is organized as follows. In section 2, I investigate putative VPE in Javanese using Goldberg's (2005) diagnostics. I show that VPE in Javanese is attested: alternative analyses are not tenable and similar characteristics to VPE in English are found. In section 3, I discuss the behavior of VPE across environments, based on shared or not shared restrictions with VP-topicalization. I argue that environments which share restrictions with VP-topicalization involve a movement account (where VP-topicalization feeds VP-ellipsis), while environments which do not involve a non-movement account. Section 4 concludes. Before moving on, I want to first point out some language traits in Javanese as well as my field methodology.

1.1. A Note on Javanese and the Field Methodology Employed

Javanese (Western Malayo-Polynesian, Austronesian) is mainly spoken on the island of Java in Indonesia. Javanese boasts three speech levels (*ngoko* 'Low', *madya* 'Mid', *krama* 'High'; e.g., Errington 1985), as well as extensive dialectal variation. The data presented here are mostly *ngoko* and based on fieldwork on dialects spoken in Semarang, Central Java; and Paciran, East Java (a village on the north coast near Tuban; see Figure 1).

³ Goldberg (2005:206): Besides English, candidates for auxiliary-stranding VPE include Moroccan Arabic, European/Brazilian Portuguese, Serbo-Croatian, and Mandarin. Sailor (2012) expands the database to include Taiwanese, Danish, and Norwegian (Standard and Northern).



Figure 1. Map of Java, Indonesia

retrieved from <http://allindonesiatravel.com/java-travel-guide/> (April 15, 2015)

Syntactic and semantic field methods were used to collect the data in this paper. Specifically, I used elicitation for grammatical as well as felicity judgments. I also asked for translations from Javanese to English or Indonesian. I worked with language consultants one-on-one as well as in groups; this way, all examples were judged by two or more people.

2. Diagnosing VP-Ellipsis in Javanese

Putative Javanese VPE must be licensed by an inflection-bearing marker, as in (1) above. Other licensors include the **auxiliaries** *wis* ‘already’, *ape* ‘FUT’, *(la)gek* ‘PROG, just’, *kudu* ‘CIRC.NEC (have to)’, *tau* ‘EXP.PRF’, *oleh* ‘DEON.POS (allow)’, *iso* ‘CIRC.POS (can)’, *gelem* ‘willing’; the **adverbs** *mesti* ‘EPIS.NEC’, *gak mungkin* ‘impossible’, *podho* ‘same’; *isek* ‘still’; **negation** *gak/ora* ‘NEG’, *durung* ‘not.yet’, *moh* ‘don’t.want’, and the **verbs** *kepingin* ‘to want’, *perlu* ‘to need’. The above markers are examples from Paciran Javanese. Counterparts that are different in Semarang Javanese are the following: *arep* ‘will, want’, *lagi* ‘PROG, just’, *enthok* ‘DEON.POS (allow)’, *oleh* ‘DEON.POS (allow)’, *isa* ‘CIRC.POS (can)’, *padha* ‘same’, *isih* ‘same’, and *ora* ‘NEG’.

Goldberg’s (2005) diagnostics for identifying VP-Ellipsis can be divided into (i) eliminating alternative analyses such as stripping and (ii) confirming VPE properties such as strict/sloppy identity. In section 2.1, I show that Javanese VPE cannot be analyzed as a different kind of ellipsis and in section 2.2, I show that Javanese VPE displays semantic and syntactic characteristics of English VPE. With these diagnostics overall, I argue that VPE is attested in Javanese.

2.1. Alternative Analyses of Putative VPE ruled out in Javanese

The possible alternative analyses are the following: (i) Stripping; (ii) Null Complement Anaphora; and (iii) Gapping plus object pro-drop. I discuss each in turn.

2.1.1. Putative VPE in Javanese Cannot be Analyzed as Stripping

The elliptical site of putative VPE in Javanese may be a larger structure (TP) as Stripping. I take Stripping to be TP-ellipsis preceded by remnant movement of the subject (and negation) to a focus position, following Depiante (2000).

To test the difference between Stripping and VPE, we can use the fact that Stripping has an ‘object’ reading in addition to a strict and sloppy reading, while VPE does not. To illustrate in English, the additional object reading for Stripping is shown in (2); this reading is not available in VPE, as shown in (3).

- (2) Pat saw herself in the mirror, and Dana too. [STRIPPING]⁴
a. strict reading: ‘...and Dana saw Pat in the mirror.’
b. object reading: ‘...and Pat saw Dana in the mirror.’
(Fortin 2007:251)
- (3) Elizabeth may go on a date with John, but Rachel may not. [VPE]
a. strict reading: ‘...but Rachel may not go on a date with John.’
b. object reading: # ‘...but Elizabeth may not go on a date with Rachel.’

Turning to Javanese, Stripping with negation has additional object reading just as in English (see (2)) as well as Indonesian (Fortin 2007):

- (4) Ali seneng mbek rawon, tapi Siti gak. [STRIPPING]
Ali happy with rawon but Siti NEG
‘Ali likes rawon, but not Siti.’
a. strict reading: ‘...but Siti doesn’t like rawon.’
b. object reading: ‘...but Ali doesn’t like Siti.’

Given that the additional object reading is available here, we can conclude that putative VPE in Javanese cannot be analyzed as Stripping.⁵

⁴ Stripping and VPE also have sloppy identity readings (ex. (2), (3), and (4)). E.g. for (2), ‘...and Dana saw herself in the mirror too.’

⁵ Stripping with negation is potentially ambiguous with VPE in Javanese, as negation also licenses VPE. Currently, only the object reading disambiguates them, but in cases where the object reading is not possible, it could be a case of VPE. Further tests are necessary to disambiguate these two cases. Thank you to Ileana Paul for pointing this out.

2.1.2. Putative VPE in Javanese Cannot be Analyzed as Null Complement Anaphora (NCA)

The elliptical site of putative VPE in Javanese may be a larger structure yet as a CP with Null Complement Anaphora (NCA). Hankamer and Sag (1976) argue that NCA targets CP and is a deep anaphor, represented as a null proform with no internal structure. This contrasts with VPE, which targets a smaller structure than TP and is a surface anaphor, represented with full syntactic structure.

To test for NCA vs. VP-ellipsis, one difference is that NCA does not allow for a ‘missing antecedent’, while VPE does. According to Hankamer and Sag (1976), this indicates that NCA does not have internal structure while VPE does.

In Javanese, the pronoun *dek-ne* ‘he/she’ following the NCA example in (5) is only variably accepted by consultants to have an antecedent referent. This contrasts with the VPE example in (6), where the pronoun is always accepted to have an antecedent referent. In other words, in the NCA example, the following pronoun sometimes has a missing antecedent, while the VPE example never has one in Javanese.

- (5) Adik-ku lali sopo sing lagek ngebel, tapi aku eling
yg.sibling-my forget who REL PROG AV.bell but 1SG remember
‘My younger brother forgot who just called, but I remember.’

%#terus dek-ne wong sugeh
then 3SG-DEF person rich
‘...and she’s rich.’

- (6) Aku durung nggolek-i konco kos, tapi Endang wes
1SG not.yet AV.look-APPL friend dorm but Endang already
‘I haven’t found a roommate yet, but Endang has.’

terus dek-ne wong-e u-apik
then 3SG-DEF person-DEF INT-good
‘...and she’s really great.’

Putative VPE in Javanese cannot therefore be analyzed as NCA given the grammaticality of (6), where the pronoun has an antecedent referent.

2.1.3. Putative VPE in Javanese Cannot be Analyzed as Gapping and Object Pro-drop

A final alternative analysis that must be considered is that the elliptical site of putative VPE in Javanese may be composed of the ellipsis/absence of more than one element, such as gapping (verb ellipsis) plus pro-drop of the object. Both

gapping and object pro-drop is available in Javanese; as shown in (7) and (8) respectively.

- (7) Arif mangan semongko, terus Tono [~~mangan~~] gedhang.
Arif AV.eat watermelon then Tono banana
'Arif ate watermelon, and Tono a banana.'
- (8) Adi durung moco buku iki, tapi Bambang wes moco [~~buku iki~~]
Adi not.yet AV.read book DEM but Bambang already AV.read
'Adi didn't read this book yet, but Bambang already read it.'

There are two different arguments to show that VPE in Javanese is not the sum of gapping plus object pro-drop. The first test is that VPE requires a licenser, while gapping does not (see (7) above). We have seen in (1) above, that putative VPE in Javanese requires an inflection-bearing licenser, such as *ape* 'FUT'. An additional example is given in (9), where the auxiliary *lagek* 'PROG, just' is obligatory when the verb phrase is elided.

- (9) Mas Adi lagek lulus SMA terus Mbak Dayu yo ***(lagek)**.
Mr. Adi just succeed high.school then Miss Dayu PRT just
'Adi just graduated high school, and then Dayu also just did.'

A second test is that VPE is compatible in syntactic islands (where the antecedent is outside the island), while gapping is not. This test shows that VPE is distinct from gapping in Javanese: ellipsis is compatible in an adjunct island for VPE as in (10), while ellipsis is incompatible in an adjunct island for gapping as in (11).

- (10) CONTEXT: *Haris belajar carane ngelangi. Haris sa'iki iso ngelangi toh gak?* 'Haris is learning how to swim. Can Haris swim now or not?'

Iyo, terus Haris ngelangi mben dino
yes then Haris AV.swim every day
sa'wise de'e wes iso [~~ngelangi~~].
after 3SG already CIRC.POS
'Yes, and Haris swims everyday now that she can.'

- (11) *Mas Joko ngeke'i gelang iku nek tante
Mas Joko AV.give-APPL bracelet DEM to aunt
sa'durunge Mas Naufal [~~ngeke'i~~] kalung iku nek tante.
before Mas Naufal necklace DEM to aunt
'J. gave a bracelet to his aunt before N. [~~gave~~] a necklace to his aunt.'

The results of these two tests provide evidence that putative VPE in Javanese cannot be analyzed as the sum of gapping and object pro-drop. Specifically, VPE in Javanese is different from gapping given the requirement of a licensing marker and given its compatibility with islands (where the antecedent is outside the island).

2.2. Putative VPE in Javanese Displays Similar Characteristics to VPE in English

In this section, I investigate whether putative VPE in Javanese has similar characteristics to VPE in English. Goldberg’s (2005) diagnostics are primarily based on English. For this reason, it is important to gain a wider database to see how robust these diagnostics are cross-linguistically and to further refine these diagnostics.

2.2.1. Strict and Sloppy Readings

A first diagnostic for VPE probes the availability of identity readings of the elided constituent. In particular, it is expected that VPE in a given language will allow both strict and sloppy readings (Goldberg 2005). A strict reading is when the referent of the pronoun in the elided constituent is identical to the referent of the pronoun of the antecedent; a sloppy reading is when the referent of pronoun in elided constituent behaves as a variable.

Turning to Javanese, putative VPE in this language allows both strict and sloppy readings, as demonstrated in (12). That is, the elided constituent *motret wong tuwo-ne* ‘photograph person old-POSS’ (where *wong tuwo* ‘person old’ is understood as the compound ‘parents’) can have the strict reading ‘...and *Joko* already took a picture of *Tomo’s* parents’. On this reading, the referent of the pronoun is identical to the referent of the pronoun in the antecedent (Tomo). The elided constituent can also have the sloppy reading ‘...and *Joko* already took a picture of *Joko’s* parents’, where the referent of the pronoun behaves as a variable, and takes the referent of the closest c-commanding antecedent; in this case *Joko*. In addition to this sloppy reading, Javanese also allows for the sloppy reading of ‘...and *Joko* already took a picture of *someone’s* parents’. This reading is due to the marker *-(n)e*, which is both a possessive marker as well as the definite marker. The referent of the elided pronoun in this case is defined by the definite marker *-(n)e*.

- (12) Mas Tomo wes [vP motret wong tuwo-ne],
 Mr. Tomo already AV.photograph person old-POSS
 terus Mas Joko yo wes [vP ~~motret wong tuwo-ne~~]
 then Mr. Joko PRT already
 ‘Tomo already took a picture of his parents, and Joko already did too.’

Strict and sloppy identity readings are therefore available in putative Javanese VPE structures.

2.2.2. No Restrictions on the Type of ‘VP’ that VPE Can Elide

A second property of VPE in English is that VPE is compatible with any internal content of the VP it elides. Goldberg uses this property as a diagnostic for VPE in other languages, as this is in contrast to Null Complement Anaphora (NCA), which only elides constituents that express a proposition.

In investigating Javanese, we find that Javanese VPE in main clauses does not impose any restrictions on the type of VP that may be elided. VPE in Javanese can elide both events and states (see (13) and (14)). Further, there are no constraints based on argument structure: Javanese VPE is compatible with intransitive verbs, transitive verbs (such as *ngulang* ‘teach’ in (13)), as well as applicatives (such as *ngeke’i* ‘to give’ in (15)).

- (13) Pak Khoim durung [_{VP} ngulang kelas biologi]. [EVENT]
 Mr. Khoim not.yet AV.teach class biology
 Pak Naufal yo durung [_{VP} ~~ngulang kelas biologi~~]
 Mr. Naufal PRT not.yet
 ‘Pak Khoim has not yet taught biology. Pak Naufal also hasn’t yet.’

- (14) Lisa gak iso [_{VP} eling embah-e], [STATE]
 Lisa NEG CIRC.POS remember grandparent-POSS
 tapi Eko iso [_{VP} ~~eling embah-e~~]
 but Eko CIRC.POS
 ‘Lisa can’t remember her grandmother, but Eko can.’

- (15) Aku mesthi [_{VP} ngeke’-i oleh-oleh nek ibuk-ku], [APPL]
 1SG EPIS.NEC AV.give-APPL RED-present to mother-my
 terus Arik yo mesthi [_{VP} ~~ngeke’-i oleh-oleh nek ibuk-ku~~]
 then Arik PRT EPIS.NEC
 ‘I’ll certainly give presents to my mother, and Arik must too.’

2.2.3. VPE Environments

Goldberg (2005) identifies five different environments where VPE may be attested in a given language: (i) syntactic islands; (ii) coordinated CPs; (iii) non-coordinated CPs (same speaker); (iv) non-coordinated CPs (different speakers; e.g. question-answer pairs); and (v) embedded clauses. I show that Javanese VPE is

attested in all of these environments, lending further support that VPE is a distinct phenomenon in Javanese.

Consider first VPE in syntactic islands, where the antecedent lies outside the island. Examples (16) and (17) demonstrate VPE within a wh-island (relative clause).

- (16) Mbak Endah gak oleh pacaran, tapi aku
Miss Endah NEG DEON.POS date but 1SG
weroh sing **oleh** [~~pacaran~~]
know REL DEON.POS
'Endah may not date, but I know who may.'

- (17) Aku ngerti ono mahasiswa sing wes moco artikel iki, tapi
1SG know exist univ.student REL already read article DEM but
aku gak iso eling mahasiswa sing ndi sing **durung**
1SG NEG CIRC.POS remember univ.student REL which REL not.yet
[~~moco artikel iki~~]
'I know some undergrads already read this article,
but I can't remember which undergrads didn't.'

(18) demonstrates VPE within a complex DP-island in Javanese:

- (18) Pak Anwar weroh laporan sing Siti wes dik promosik-no
Pak Anwar know report REL Siti already PASS promote-APPL
mbek laporan sing Kana **durung** [~~dik promosik-no~~]
with report REL Kana not.yet
'Pak Anwar saw the report that Siti already was promoted, and the report
that Kana wasn't yet.'

VPE in Javanese is also compatible in adjunct islands, as already shown in (10) above. An additional example is given in (19):

- (19) Agung ngeresik-i jedhing mergo-ne Ipul **gak** [~~ngeresik-i~~]
Agung AV.clean-APPL bathroom because-NE Ipul NEG
'Agung cleaned the bathroom because Ipul didn't.'

A second environment where VPE is attested in Javanese is in coordinated CPs. In both Semarang and Paciran Javanese varieties, VPE is acceptable with *tapi* 'but' and *terus* '(and) then', as shown in (20) and (21). VPE is also accepted with *lan* 'and', but only in Semarang Javanese.

- (20) Mbak Ayu oleh pacaran, **terus** Mbak Putri yo oleh [**pacaran**].
Miss Ayu DEON.POS date then Miss Putri PRTDEON.POS
'Ayu was allowed to date, and then Putri was also allowed to.'
- (21) Mbak Ayu gak oleh pacaran, **tapi** Mbak Putri oleh [**pacaran**].
Miss Ayu NEG DEON.POS date but Miss Putri DEON.POS
'Ayu is not allowed to date, but Putri is allowed to.'

Non-coordinated CPs (uttered by the same speaker) is a third environment where VPE is found in Javanese, as illustrated in (22) and (23).

- (22) Mas Adi lagek lulus SMA.
Mr. Adi just succeed high.school
Mbak Dayu yo lagek [~~lulus~~—SMA].
Miss Dayu PRT just
'Adi just graduated high school. Dayu also just did.'
- (23) Mas Doni iso ngelangi. Mas Yoko yo iso [**ngelangi**].
Mr. Doni CIRC.POS AV.swim Mr. Yoko PRT CIRC.POS
'Doni can swim. Yoko can also swim.'

The fourth environment that Goldberg (2005) notes for VPE in a given language is non-coordinated CPs with two different speakers, such as question-answer pairs. VPE in Javanese is also compatible in this environment: (24) illustrates VPE as an answer to a yes-no question, and (25) illustrates VPE as an answer to a wh-subject question.

- (24) Q: mbak Nunung tau lungo reng Jakarta toh?
Miss Nunung EXP.PRF go at Jakarta FOC
'Has Miss Nunung ever gone to Jakarta?'

A: ? Iyo, Nunung tau.
yes Nunung EXP.PRF
'Yes, Nunung has.' (Vander Kloek 2012:166)

- (25) Q: Sopo sing iso nyopir sepeda montor? A: Aku iso.
who REL CIRC.POS AV.drive bike motor 1SG CIRC.POS
'Who can drive a motorbike?' 'I can.'

Embedded clauses is the fifth environment where a language is expected to allow VPE. This environment is also borne out for Javanese, as shown in (26).

- (26) Tak pikir Pak Abbas iso ngomong boso Sunda,
1SG.CL think Mr. Abbas CIRC.POS AV.speak language Sunda
tapi Pak Agus mikir nek de'e gak iso.
but Mr. Agus AV.think COMPL 3SG NEG CIRC.POS
'I think that Mr. Abbas can speak Sundanese, but Mr. Agus thinks that he cannot.'

These examples show that Javanese VPE occurs in all of the environments that English VPE occurs in—syntactic islands; coordinated CPs; non-coordinated CPs (same speaker); non-coordinated CPs (different speakers; e.g. question-answer pairs); and embedded clauses—giving further support that VPE is indeed attested.

2.3. Summary: VPE is Attested in Javanese

Based on Goldberg's (2005) diagnostics, we can conclude that VPE is attested in Javanese. I have shown that Javanese VPE cannot be analyzed as a different kind of ellipsis such as Stripping, Null Complement Anaphora, or the sum of gapping plus object pro-drop. I have also shown that VPE in Javanese displays semantic and syntactic characteristics of English VPE. These properties include the availability of both strict and sloppy identity readings, the lack of restrictions on the type of *vP* it can elide, and the ability to occur in all environments as other auxiliary-stranding VPE languages.

3. Towards an Analysis of Javanese VPE

In this section, I discuss the type of analysis that is most appropriate for Javanese VPE. I argue that Javanese VPE in fact must involve two different strategies, which are dependent on the environment. I provide arguments that VPE in syntactic islands, coordinated CPs, and non-coordinated CPs uttered by the same speaker must involve a non-movement analysis, while VPE in non-coordinated CPs uttered by different speakers and in embedded clauses must involve a movement analysis, where VP-topicalization feeds VP-ellipsis (e.g. Johnson 2001 for English; Szczegielniak 2006 for Polish and Russian). Interestingly, this partition is different from closely related Indonesian, where VPE is attested but only a non-movement analysis is proposed (Fortin 2007).

3.1. Non-movement Analysis: VPE in Islands, Coordinated CPs, and Non-coordinated CPs (Same Speaker)

First, I argue that Javanese VPE in islands, coordinated CPs, and non-coordinated CPs (same speaker) must involve a non-movement analysis. By non-movement, I mean that VPE cannot be derived via VP-topicalization (see (27)), followed by PF-deletion of the VP in topic position (e.g. Johnson 2001; Szczegielniak 2006).

- (27) VP-topicalization: *movement of the ‘VP’ to Spec, TopicP*
 [TopicP **XP** ...[TP SUBJ... [... [**txP**]]]]

Instead, I suggest that an Agree-type analysis such as in Aelbrecht (2010) is most appropriate for Javanese VPE in these environments. Evidence that a movement account is untenable for these environments is that VPE has no restrictions, while VP-topicalization has two constraints.

The first restriction is that VP-topicalization is only possible with events (Vander Klok and Déchaine 2014).⁶ Compatibility of VP-topicalization with events is demonstrated in (28); while the incompatibility with states is shown in (29).

- (28) **nggotong watu-ne**, cak Kholiq iso ✓EVENT
 AV.lift rock-DEF Mr. Kholiq CIRC.POS
 ‘Lift the stone, Kholiq can.’ (Vander Klok 2012:153)

- (29) CONTEXT: *Opo Mas Adi iso eling ceritone mbahe Adi?* *STATE
 ‘Could Mas Adi remember his grandfather’s story?’

***eling cerito-ne mbah-e** Adi, mas Adi iso
 remember story-DEF grandparent-DEF Adi Mr. Adi CIRC.POS
 [‘Remember the story of Adi’s grandfather, Mr. Adi could.’]
 (Vander Klok and Déchaine 2014:558)

In contrast, VPE does not have this restriction for islands, coordinated CPs, and non-coordinated CPs with the same speaker. That is, VPE in these environments is possible with both events and states. (30) and (31) shows the compatibility of VPE in coordinated CPs with both events and states respectively.

- (30) Aku wis **mangan** terus Fina barang wis. ✓EVENT
 1SG already AV.eat then Fina together already
 ‘I already ate, and then Fina also finished.’

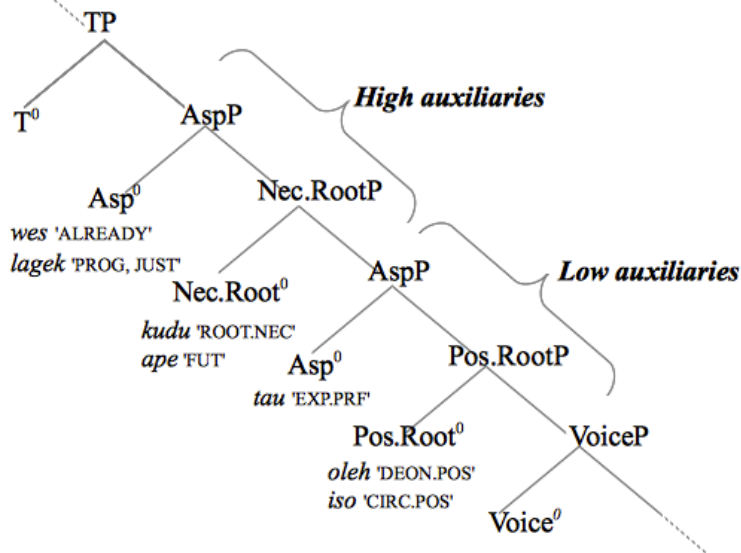
- (31) Lisa gak iso **eling mbah-e**, tapi Eko iso. ✓STATE
 Lisa NEG CIRC.POS remember grandparent-DEF but Eko CIRC.POS
 ‘Lisa can’t remember her grandfather, but Eko can.’

Given that VPE does not have the same event-type restriction (*state) as VP-topicalization, this suggests that VPT does not feed VPE as part of its derivation for these environments.

⁶ See Vander Klok and Déchaine (2014) for tests (based on both cross-linguistic and language-specific diagnostics) that distinguish events from states in Javanese.

The second restriction for VP-topicalization in Javanese is that it is licensed only by a subset of auxiliaries, which correspond to a class of syntactically ‘low’ auxiliaries (Vander Klok 2012). Other auxiliaries, which are syntactically high, do not license VP-topicalization. Despite extensive dialectal variation in Javanese, this partition is a language-wide phenomenon (Vander Klok 2015). To illustrate in Paciran Javanese, the low auxiliaries which license VP-topicalization include *tau* ‘EXP.PRF’, *oleh* ‘DEON.POS’, *iso* ‘CIRC.POS’, and *gelem* ‘willing’, while the high auxiliaries which do not license VP-topicalization include *wes* ‘already’, *lagek* ‘PROG, just’, *ape* ‘FUT’, *kudu* ‘ROOT.NEC’:

- (32) Two syntactic classes of auxiliaries in Paciran Javanese (Vander Klok 2012; 2015)



This restriction for VP-topicalization in Javanese is demonstrated in (33); *lagek* ‘PROG’, as representative of the class of high auxiliaries, cannot license VP-topicalization; while *iso* ‘CIRC.POS’, as representative of the class of low auxiliaries, can.

- (33)a. *tuku beras, pak Suwanan **lagek** *HIGH AUX
 buy uncooked.rice Mr. Suwanan PROG
 (‘Buying rice, Pak Suwanan is just.’)
- b. nggotong watu-ne, cak Kholiq **iso** ✓LOW AUX
 AV.lift rock-DEF Mr. Kholiq CIRC.POS
 ‘Lift the stone, Kholiq can.’

In comparison, VPE in syntactic islands, coordinated CPs, and non-coordinated CPs with the same speaker is licensed by *all* auxiliaries in Javanese. (34) illustrates

that both high and low auxiliaries license VPE with non-coordinated CPs with the same speaker. A number of examples above show that high auxiliaries can license VPE with coordinated CPs as well: see (1), (6), (9), (12), and (30).

- (34)a. Mas Adi lagek lulus SMA. Mbak Dayu yo **lagek**.
Mr. Adi just succeed high.school Miss Dayu PRT just
'Adi just graduated high school. Dayu also just did.' ✓HIGH AUX
- b. Mas Doni iso ngelangi. Mas Yoko yo **iso**.
Mr. Doni CIRC.POS AV.swim Mr. Yoko PRT CIRC.POS
'Doni can swim. Yoko can also swim.' ✓LOW AUX

Since VPE in islands, coordinated CPs, and non-coordinated CPs with the same speaker does not have the same restriction for high auxiliaries as VP-topicalization, it suggests that VP-topicalization in these environments does not feed VPE as part of its derivation.

I propose that VPE in these environments are best analyzed with an Agree-type analysis, such as Aelbrecht (2010). Aelbrecht (2010) builds on Merchant (2001) and proposes that ellipsis is licensed via Agree between the licensing head and an E-feature on the deleted constituent (marked for deletion at PF). In Javanese, the licensing head can involve Asp⁰, Nec.Root⁰, Pos.Root⁰, or Neg⁰ as shown in (32) above.⁷ Crucially, in this analysis, no movement is involved in the overt syntax, capturing the fact that the VP-topicalization restrictions are not shared by VPE in these environments.⁸

3.2. Movement Analysis: VPE in Non-coordinated CPs (Different Speaker) and Embedded Clauses

In non-coordinated CPs with different speakers (henceforth question-answer pairs; both yes-no and wh-subject questions) and embedded clauses, VPE behaves parallel to VP-topicalization with respect to the two restrictions outlined above. That is, VPE in question-answer pairs and embedded clauses are restricted to events and low auxiliaries in Javanese.

First, to illustrate the event restriction with VPE in question-answer pairs, (35) shows that VPE in an answer to a yes-no question is acceptable when the elided material is an event, such as *ngelangi* 'to swim'. However, VPE in an answer to a yes-no question is not possible when the elided material is a state, as shown in (36) with *ngerti* 'to understand'.

⁷ It may be that these markers move to T⁰ to license VP-ellipsis in Javanese as typically T⁰ is argued to be the licensing head (see e.g. Lobeck 1995, Aelbrecht 2010). However, concrete evidence is necessary to validate this possibility.

⁸ Other analyses may be possible, such as an LF-Copying approach (movement in covert syntax) as suggested in Fortin (2007) for VPE in Indonesian.

Similarly, VPE within embedded clauses are ungrammatical with high auxiliaries, but grammatical with low ones, as illustrated with *lagek* ‘PROG, just’ and *iso* ‘CIRC.POS’ in (38).

- (38) a. Tak pikir mbak Lisa lagek mepe klambi nek dhuwur,
 1SG.CL think Miss Lisa PROG AV.hang clothes at above
 * terus adik-ku mikir mbak Lisa yo **lagek**. *HIGH AUX
 then yg.sibling-my AV.think Miss Lisa PRT PROG
 (‘I think that Lisa is hanging the clothes upstairs, and my younger sister
 also thinks that Lisa is doing so.’)
- b. Tak pikir Pak Abbas iso ngomong boso Sunda,
 1SG.CL think Mr. Abbas CIRC.POS AV.speak language Sunda
 tapi Pak Agus mikir nek de’e gak **iso**. ✓LOW AUX
 but Mr. Agus AV.think COMPL 3SG NEG CIRC.POS
 ‘I think that Mr. Abbas can speak Sundanese, but Mr. Agus thinks that he
 cannot.’

That VPE in embedded clauses and question-answer pairs have the same restrictions based on the syntactic class of auxiliary as VP-topicalization is evidence for a movement account. Specifically, these data suggest that VP-topicalization feeds VPE as part of its derivation, as argued for verb-stranding verb ellipsis in Polish and Russian (Szczegielniak 2006) and proposed for auxiliary stranding VPE in English (Johnson 2001).

I propose that for Javanese VPE in these environments, the constituent to be elided first topicalizes (movement to Spec, TopP in the left periphery, above TP) in the overt syntax, and then is deleted at PF, as outlined in (39).

- (39) VP-topicalization followed by PF-deletion:
- a. $[_{\text{TopicP}} \mathbf{VP} \dots [_{\text{TP}} \text{SUBJ} \dots \text{T} [\dots [\mathbf{t}_{\text{VP}}]]]]]$ *Movement in overt syntax*
- b. $[_{\text{TopicP}} \mathbf{VP} \dots [_{\text{TP}} \text{SUBJ} \dots \text{T} [\dots [\mathbf{t}_{\text{VP}}]]]]]$ *Deletion at PF*

The event and auxiliary restrictions on VPE in these environments are accounted for by the same analysis proposed for VP-topicalization in Javanese. That is, these restrictions are argued to be governed by independent syntactic principles; the interaction of phasal properties and anti-locality (cf. Abels 2003; Grohmann 2003). The general outline for when ungrammaticality results (with a high auxiliary or a state) is that movement of a complement XP under a phase head is impossible: movement out of a phase must go through the specifier of the phase but that very movement is too short (anti-locality). Vander Kloek (2012) proposes a phase head above vP but below TP to capture the ungrammaticality/grammaticality of high and low auxiliaries in VP-topicalization in Javanese. Vander Kloek and Déchaine

(2014) account for the event/state restriction in VP-topicalization under the vP phase in Javanese (e.g. Sato 2010, 2012).

While I leave some details for the VPE analysis via movement open at this point (such as how the XP is marked for deletion; e.g. E-feature, Merchant 2001); importantly, this account will capture the fact that only VPE in embedded clauses and question-answer pairs share the restrictions of VP-topicalization.

4. Conclusion

In this paper, I have argued that VPE is attested in Javanese: it cannot be analyzed alternatively and it shares syntactic and semantic characteristics with established VPE. Javanese adds to the database of languages with VPE, and shows that Goldberg's (2005) diagnostics for identifying VPE are cross-linguistically robust.

Further, a novel derivational partition is revealed in the different environments for VPE. Specific environments in Javanese are suggested to have different derivations based on parallel restrictions with VP-topicalization or not. I propose that VPE in syntactic islands; coordinated CPs; and non-coordinated CPs (same speaker) requires a non-movement account since these do not share restrictions with VP-topicalization, while VPE in question-answer pairs and embedded clauses requires a movement account (where VP-topicalization feeds VPE) since these do share the same constraints as VP-topicalization.

References

- Abels, Klaus. 2003. Successive Cyclicity, Anti-locality and Adposition Stranding., University of Connecticut Ph.D. dissertation.
- Aelbrecht, Lobke. 2010. *The syntactic licensing of ellipsis*. Amsterdam, New York: John Benjamins.
- Depiante, Marcela. 2000. The syntax of deep and surface anaphora: a study of null complement anaphora and stripping/bare argument ellipsis, University of Connecticut: Ph.D. dissertation.
- Fortin, Catherine. 2007. Indonesian sluicing and verb phrase ellipsis: Description and explanation in a minimalist framework, Linguistics, University of Michigan: PhD.
- Goldberg, Lotus. 2005. Verb-Stranding VP-Ellipsis: A cross-linguistic survey, Linguistics, McGill University: PhD.
- Grohmann, Kleanthes. 2003. *Prolific Domains*. Amsterdam: John Benjamins.
- Johnson, Kyle. 2001. What VP Ellipsis can do, what it can't, but not why. In *The handbook of contemporary syntactic theory*, eds. Mark Baltin and Chris Collins, 439-479. Oxford, Boston: Blackwell Publishers.
- Lobeck, Anne. 1995. *Ellipsis: Functional heads, licensing, and identification*. New York: Oxford University Press.
- Sailor, Craig. 2012. Tag questions and the typology of VP-ellipsis. Ms. UCLA.

The Proceedings of AFLA 22

- Sato, Yosuke. 2010. *Minimalist interfaces : evidence from Indonesian and Javanese*. Linguistik aktuell = Linguistics today. Amsterdam ; Philadelphia: John Benjamins Pub. Co.
- Sato, Yosuke. 2012. Successive cyclicity at the syntax-morphology interface: Evidence from Standard Indonesian and Kendal Javanese. *Studia Linguistica* 66:32-57.
- Sato, Yosuke. 2013. Cyclic spell-out and modal complement ellipsis in Javanese. *Linguistic Analysis* 38:183-206.
- Szczegielniak, Adam. 2006. VP-ellipsis and VP-topicalization. In *NELS 35 Proceedings*, eds. Leah Bateman and Cherlon Ussery, 603-615: GLSA UMass, BookSurge Publishing.
- Vander Klok, Jozina, and Déchaine, Rose-Marie. 2014. Stative versus eventive predicates and vP-internal structure. In *Proceedings of Berkeley Linguistics Society 40*, eds. Herman Leung et al., 544-578. Berkeley: BLS.
- Vander Klok, Jozina. 2012. Tense, aspect, and modality in Paciran Javanese, McGill University: Ph.D. dissertation.
- Vander Klok, Jozina. 2015. The dichotomy of auxiliaries in Javanese: Evidence from two dialects. *Australian Journal of Linguistics* 35.2: 142-167.

THE INDONESIAN DP

Ekarina Winarto
Cornell University
ew385@cornell.edu

In this paper¹ I explore the constituent order within the Indonesian DP and argue that Indonesian employs two kinds of definite articles : (i) an anaphoric article, and (ii) a uniqueness article, which have different semantic properties (along the lines of Schwarz 2009) and are encoded in two different DP-layers. My proposal has interesting typological implications, since it suggests that in addition to NP and DP languages (e.g., Bošković (2008), Despic (2011), there are also languages with a layered DP structure. I also argue that the uniqueness DP layer and ClassifierP are in fact one and the same phrase. This is supported by the fact that the uniqueness definite article and classifiers are in complementary distribution; that classifiers are closely related to the uniqueness DP layer has also been suggested elsewhere. My proposal also derives in a natural way the order of DP-internal elements in Indonesian, and provides further support for Cinque's (2005) roll-up theory of the DP.

1. Introduction

Indonesian, the language that I will be exploring in this paper has an overt definite article in the form of a suffix *-nya*. Many scholars like Abney (1987) have suggested, on the basis of English data, that the definite article, indefinite article, possessives and demonstratives are all the head of a single DP projection, collectively categorized under the umbrella term Determiner (D). This theory has been challenged by many other scholars on the basis of various data from languages other than English where these types of determiners can co-occur with each other (Cinque (2005) is an example). Even scholars working with English have also posited more than one DP projection category for the language.

Heycock and Zamparelli (2003, 2005) have posited 3 DP layers: DP, NumberP and PIP. In their analysis, DP is the layer that contains the strong determiners universally in all languages; these are elements like the definite article, demonstratives and quantifiers. NumberP is the locus of cardinality and contains numerals, as well as weak determiners like the indefinite article. Lastly, the PIP is the projection that is responsible for plurality. In their analysis, definite articles can occur both in PIP and in DP; while the indefinite article can only occur in the PIP.

¹ I would like to thank my Q-paper chair, Miloje Despic, who has helped me develop the ideas and concepts discussed in this paper; also I would like to thank my A-paper chair, John Whitman, who has proof-read and helped me in writing this paper. Also equally helpful in the development of the ideas here are my committee members John Bowers, Molly Diesing and Abby Cohn.

Borer (2005), on the other hand, claims that the 3 maximal projections needed in a layered DP should be the DP, #P and ClassP. In her analysis, she claims that all nouns are mass, and that they need to be portioned out before they interact with the counting system. This portioning out is done by the classifier in classifier languages like Mandarin, but through plural morphology in languages like English. #P is the quantity phrase for Borer; and the indefinite article can be either in ClassP or #P; the definite article can be associated with all three projections.

These arguments for the existence of a universal layered DP are mainly based on the idea that different semantic meanings should be translated into different maximal projections in the syntactic structure. Ihsane and Puscas (2001), for example, are among the scholars who have argued that definiteness can be divided into specificity and definiteness, and that these semantic differences should be encoded into the syntax in the form of a split/layered DP, with each layer having distinct semantic content. More recently, Schwarz (2009) has also claimed that there are two kinds of definite articles in German; he extends this to a cross-linguistic claim in his subsequent publications, although he has not made any claims about how the semantics of definiteness should be translated into the syntactic structure.

In Indonesian, there is also evidence for two types of definite articles, one denoting the absolute or situational uniqueness of its referent, and another one denoting definiteness of its referent by way of strict co-indexation of repeated occurrences of the same noun within the same discourse. Sometimes the use of both kinds of definite article can overlap in Indonesian, but there are cases where one type of article is at least preferred to the other. Interestingly, the two kinds of definite article can also co-occur.

I will give a brief description of the Indonesian NP data in the next subsection. Then, in section 2 of the paper, I will expand on the use and co-occurrence possibilities of both kinds of definite article, as well as explain in more detail the structure of the layered DP that I am proposing. In section 3 of the paper, I will then discuss the co-occurrence impossibility of the article *-nya* and the number (and classifier). Then in section 4, I will provide some evidence supporting Cinque's (2005) claim that various attested surface DP orders in natural languages are the result of phrasal movements.

1.2. The Indonesian DP in a nutshell

The general word order in the Indonesian DP is such that all noun modifiers like adjectives and determiners are placed post-nominally, with the exception of number and classifier. The ordering of the post-nominal elements is the opposite of the DP word order in English.

The number of direct attributive adjectives that can be placed post-nominally is limited to a maximum of two and the order of post-nominal adjectives in Indonesian is the mirror order of prenominal adjectives in English, similar to

what others have observed in Romance languages and other languages with post-nominal adjectives.

- (1) Bola merah besar
ball red big
'big red ball'
- (2) * merah bola besar
red ball big
- (3) ? Bola besar merah
ball big red

This limitation to two postnominal adjectives only applies to direct attributive adjectives; predicative adjectives contained within a relative clause can be freely added with no restriction on their number. Relative clauses in Indonesian are headed by *yang*, a close equivalent to the English *that*.

- (4) ?Bola merah besar baru
ball red big new
'new red big ball'
- (5) Bola merah besar yang baru
ball red big that new
'the big red ball that is new'
- (6) Bola yang merah besar dan baru
ball that red big and new
'the ball that is big, red and new.'

From the contrast we see in (4),(5) and (6), it is clear that there is a difference between attributive and predicative adjectives. While attributive adjectives can directly modify the noun, predicative adjectives can only modify the noun indirectly from within a relative clause.

Possessive forms in Indonesian are identical to the pronouns: they are placed postnominally just like the adjectives. Possessives can co-occur together with the adjectives, and are placed after the adjectives.

- (7) Bola merah saya
ball red my
'my red ball'
- (8) Bola saya merah
ball my red
'my ball is red'

Even though most of the possessive forms in Indonesian are identical to their pronominal counterparts, there exists a special form of the third person singular possessive: *-nya*. The pronominal form of the third person singular is *dia*,

and when placed behind a noun, *dia* can also signal the third person singular possessive. However, as an alternative, *-nya* can also be used.

- (9) Bola merah dia
ball red his/her
'His or her red ball.'
- (10) Bola merah-nya
ball red his/her
'His or her red ball / the red ball'

There is no contrast between (9) or (10) in terms of signaling possession. However, there is ambiguity present in sentence (10) because *-nya* also happens to be the definite article in Indonesian. As such, sentence (10) can be interpreted as either 'his/her red ball' or 'the red ball' depending on the discourse. *-nya* cannot co-occur with other possessives, even when it is used as the definite article.

Then, there are two kinds of what people generally call "demonstratives" in Indonesian: *ini* 'this' and *itu* 'that'. The demonstratives can co-occur with the adjectives and the possessives; and are placed after the possessives.

- (11) Bola merah saya ini
ball red my this
'this my red ball'
- (12) Bola merah saya itu
ball red my that
'that my red ball'

Aside from serving as a demonstrative, *itu* can also be an anaphoric definite article. If a noun has been linguistically introduced in the discourse, then *itu* can be used as a definite article that refers to that specific noun.

- (13) Saya baru beli baju baru. Baju itu mahal sekali.
I just buy clothes new. Clothes DEF expensive very
'I just bought some new clothes. The clothes are very expensive.'

Although Austronesian languages typically have a vast variety of classifiers, the use of these classifiers has become very sparse in Indonesian especially in colloquial/everyday speech. Classifiers are often omitted with numerals that denote two or more, but they are still used very frequently when denoting the number one or when used with uncountable nouns.

When numbers and classifiers co-occur together, both are placed in front of the noun with the classifier placed between the number and the noun. Both numbers and classifiers can co-occur with all other elements in the DP except for the definite article *-nya*.

- (14) Tiga (buah) bola merah saya itu.
three class. ball red my that
'Those three balls of mine'
- (15) *satu (buah) matahari -nya
one class. sun DEF
'the one sun'
- (16) * Dua orang presiden -nya baru mendarat di Soekarno-Hatta.
two class. president DEF just landed at Soekarno-Hatta.
'The two presidents just landed at Soekarno-Hatta (the Jakarta airport)'
- (17) *Lima (buah) buku -nya mahal sekali
five class. book DEF expensive very
'The five books are very expensive.'

Even though the classifiers are not necessarily always pronounced, nothing can intervene between the number and the classifier, and not pronouncing the classifier does not change the order or co-occurrence possibilities of the number with other elements in the DP. The presence of the classifier *buah* in sentence (14) or (15), for example, does not really matter. As long as the numeral is present, *-nya* cannot be used. Thus, I conclude that even if the classifier is not pronounced at times, it is still present in the syntactic representation.

Also, to form the indefinite article, a reduced form of the numeral *satu* 'one', *se-* is used together with a classifier.

- (18) Saya baru membeli sebuah buku.
I just buy one-classifier book
'I just bought a book.'

Very often, the full numeral *satu* 'one' can also be used to replace the reduced form of one and classifier to form the indefinite article; in which case *satu* functions simply as an indefinite article on its own.

- (18') Saya baru membeli satu buku.
I just buy one book
'I just bought a book.'

Due to this close relationship between the number and classifier, I propose in this paper that the number and classifier are contained within one maximal projection.

The following points summarize the data in relation to the determiners:

- A. The demonstrative and definite article CAN co-occur.
- B. The possessives and the demonstrative CAN co-occur.
- C. The indefinite article and the possessives CAN co-occur.
- D. The indefinite article and the demonstrative CAN co-occur.
- E. The definite article *-nya* and possessives CANNOT co-occur.

- F. The definite article and indefinite article (as well as other Number+Classifier combination) CANNOT co-occur.

At first glance, it seems that there is a need for more than just one projection to “house” the determiner, but at the same time an overgeneration of maximal projections in the DP is also not desirable. Thus, in the following sections, I explore this data more in depth, and will propose which maximal projections are needed within the Indonesian DP.

1.3. My proposal in a nutshell

Considering the co-occurrence possibilities of the different determiners in Indonesian, I propose a need for 3 layers of DP: DP1, DP2 and the Possessive Phrase. I argue that there are two kinds of definiteness that can be conveyed in slightly different ways in Indonesian: strong definiteness (based on anaphoricity) and weak definiteness (based on situation). These different kinds of definiteness provide the basis for a layered DP in the syntactic structure. In my proposal, the DP layer with strong definiteness is the topmost layer DP1, and houses the strong article as well as demonstratives. The DP layer with the weak definiteness is DP2² and can be filled with either the weak article *-nya*, the indefinite article or the Number and Classifier. Based on this data and analysis, specifically based on the fact that elements occupying DP2 can surface either to the right or to the left of the Noun, I argue that Cinque’s rather complicated roll-up movement provides the best account of the surface order in the Indonesian DP.

2. Definite Articles & Demonstratives

2.1. Two Kinds of Definite Articles in Indonesian

The definite article *itu* is the preferred definite article for indicating anaphoricity, while *-nya* is the only article for definite descriptions that are used situationally and in bridging. Sentence (19) below (repeated from sentence (13) above) is an example of how *itu* is used anaphorically as a definite article.

- (19) Saya baru beli **baju baru**. **Baju itu** mahal sekali.
I just buy clothes new. Clothes DEF expensive very
‘I just bought some new clothes. The clothes are very expensive.’

Here, *itu baju* clearly denotes the *baju* that has already been concretely established in the course of this specific discourse. *-nya* can also be used anaphorically (sentence (19’) below), but *itu* is preferred. This is not surprising,

² This is a temporary naming that I have adopted. This layer can also simply be the NumP or ClassP layer, as has been suggested by many others and has been pointed out to me by a member of my Q paper committee, John Bowers.

because when an entity has been established in the discourse, then that entity also becomes the most salient entity in that particular situation; thus, it is expected that *-nya* should also be felicitous. This means that even though, *-nya* is felicitous in sentence (19') below, there is a different mechanism at play compared to sentence (19) above. I propose that in sentence (19) above, there is strict co-indexation of both occurrences of *clothes*, but in sentence (19') below the definite description picks up its referent because it is sufficiently "unique" in the discourse due to previous mention in the preceding sentence. In this case, definiteness is derived without co-indexation.

- (19') Saya baru beli baju baru. Baju -nya mahal sekali.
I just buy clothes new. Clothes DEF expensive very
'I just bought some new clothes. The clothes are very expensive.'

However, if a definite description refers back anaphorically to a preceding entity that has a different form in the utterance, only *itu* is felicitous.

- (20) Saya baru saja mengunjungi seorang **anggota DPR**.
I just visited a member parliament.
Politikus itu (?nya) sangat ramah.
politician DEF very friendly.
'I just visited a member of parliament. The politician was very friendly.'

Here, *the politician* refers back to *the member of parliament* that has been uttered in the discourse; and for this specific anaphoric use, only *itu* is grammatical. This is the case because *itu*, as I proposed above, picks up its referent by way of co-indexation (in this case between *member of parliament* and *politician*). *-nya*, on the other hand, must pick its referent by virtue of it being "unique" in the situation. Because there are a lot of members of parliament, and there are several types of politician in the world, this referent (*anggota DPR*) is not unique enough to be picked up by *-nya*. This shows that there is indeed a contrast between how *itu* and *-nya* pick up their referent in definite descriptions.

To indicate absolutely unique entities like *the sun*, however, the definite article *-nya* is used exclusively.

- (21) Matahari-nya(*itu) panas sekali hari ini.
sun DEF hot very day this
'the sun is very hot today.'

And, if there is only one table in a room where a conversation takes place, then when referring to that specific table in the room, the discourse participant would also use *-nya* (assuming that *the table* has never been verbally mentioned in the conversation). This is expected because in sentences (21)-(22), coindexation is not possible and *-nya* requires some sort of uniqueness entailment.

- (22) Mejanya (*itu) besar sekali
table DEF big very
'The table is very big.'

For bridging uses, only *nya* can be used (See sentence (23) below).

- (23) Budi baru membeli **sepasang sepatu. Designer-nya (*itu)** terkenal di Paris.
Budi just buy a pair shoes designer DEF famous in Paris.
'Budi just bought a pair of shoes. The designer is famous in Paris.'

From the sentences above, it is clear that *itu* has a more "anaphoric" use where it can only describe the definiteness of an entity that has been linguistically introduced in a specific conversation (indicating strong familiarity and requiring strict coindexation)³. On the other hand, *-nya* can pick out a unique entity in a particular situation, even if this entity has not been concretely introduced in the conversation (only requiring weak familiarity). As demonstrated in sentences (13) and (13'), both types of definite articles can overlap in their use, but for strictly situational use (as demonstrated in sentences (20) and (21)), *itu* cannot be used.

Schwarz (2009), in his dissertation also argues that there are two kinds of definiteness in German. I reproduce the data below taken from his dissertation to show some parallelism with the Indonesian data:

- (24) Hans ging zum Haus.
Hans went to-the weak house
'Hans went to the house.'
- (25) Hans ging zu dem Haus.
Hans went to the strong house
'Hans went to the house.'

In Schwarz's theory, the contracted form of the definite article (in (24)) is called the weak article; it is characterized as requiring uniqueness (relativized to a situation). The non-contracted form of the article (in (25)), on the other hand, is the strong article, which is used anaphorically. When we consider how strong definiteness is realized, it makes sense to have this clear division. Used anaphorically, an article signals strong definiteness, where the definiteness does not depend on a situation. Revisiting my sentence (13), for example, the identity of the clothes that the speaker talks about is fixed; and even if she/he buys many more clothes in the future, or conducts the conversation on Mars, the definiteness of the clothes does not change because all that matters is the coindexation of both instances of *clothes* in the utterance. The definiteness indicated in sentence (21) above, however, may change depending on the context. If, one were to imagine the

³ At this point, I cannot explain why nested coindexation is not allowed in sentence (23).

conversation taking place in another galaxy (not the Milky Way), then *the sun* that the speaker refers to might change.

Interestingly, *itu* and *-nya* can also both co-occur as definite articles in Indonesian. However, when *itu* and *-nya* co-occur, for both elements to be interpreted as the definite articles, the context must be clear. If we take sentence (26), for example, I would first have to have already introduced a bicycle in the discourse to make the sentence felicitous (shown in sentence (26')).

- (26) ?Sepedanya itu baru dicuri orang
bicycle DEF DEF just stolen person
'The bicycle just got stolen by someone.'
- (26') Ali baru beli sepeda baru, tapi sepeda(-nya) itu sudah dicuri orang
Ali just buy bicycle new but bicycle DEF DEF already stolen person
kemarin.
yesterday.
'Ali just bought a new bicycle, but that bicycle has been stolen by someone
yesterday.'

When there is double occurrence of the two different kinds of definite article, the meaning of the noun phrase becomes more specific. With *-nya* added to *itu*, as seen in sentence (26') above, the speaker isn't just picking out a specific bicycle, but he/she seems to also be familiar (either has seen, or has ridden, etc) with the bicycle. This means that definite article doubling is not merely an instance of morphological doubling, and supports the argument that both kinds of definite article have their own semantic meanings that are different. *Itu* picks out the previously mentioned bicycle via coindexation, while *-nya* also picks out the very same bicycle by way of it being unique (from the standpoint of speaker knowledge). Thus, there is the added layer of definiteness present in the sentence.

From the data in this section, it seems that there are these two kinds of definiteness that, in Indonesian, can be expressed by two different definite articles: i) *itu* is used more for anaphoricity (strong definiteness) and must derive definiteness from strict coindexation; while ii) *-nya* is used more for situational and bridging purposes. Ihsane and Puscas (2001) and Alexiadou (2014), among others, have suggested that in cases where there are different kinds of definiteness, as we have seen is the case in Indonesian, the DP should be split into different categories to represent the different semantic meanings. Following this line of analysis, I posit two layers of DP. I propose calling these two layers DP1 and DP2 (See Figure 1 below). DP1 represents strong definiteness; and DP2 represents weak definiteness.

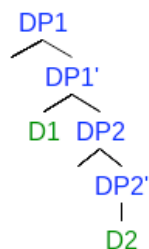


Figure 1: A layered DP

This split in the DP to house strong and weak articles is not a novel idea. Borer (2005), Heycock & Zamparelli (2003), and others, have also proposed that the strong article like the definite article should be housed in a different maximal projection than the weak articles like the indefinite article.

2.2. Demonstratives

Now, looking at the function of the demonstratives in Indonesian, I hesitate to group the two demonstratives *ini* ‘this’ and *itu* ‘that’ into the same category. If they are both demonstratives, then we would expect them to function similarly as regards their usage as an anaphoric definite article. However, this is not the case in Indonesian: *ini* cannot be used anaphorically in the language.

- (27) Saya baru beli baju baru di Mangga dua. Baju baru itu (?ini)
 I just buy clothes new in Mangga dua. Clothes new DEF (?this)
 murah sekali
 cheap very
 ‘I just bought some clothes in Mangga dua (a famous retail shopping place in Jakarta). The clothes are very cheap.’

From this, it has to be concluded that *ini* is a purely demonstrative element, whereas *itu* has an additional meaning in its semantics that allows it to be an anaphoric definite article. Also, *ini* can only be used demonstratively to point to the most recently mentioned entity in a preceding sentence.

- (28) Dono punya seekor kucing putih dan seekor kucing hitam. Kucing hitam
 Dono has a cat white and a cat black cat black
 ini lebih lama dipelihara Dono dibandingkan kucing-kucing lainnya.
 ini more long taken care Dono compared cats other
 ‘Dono has a white cat and a black cat. Dono has taken care of this black cat for a longer time compared to his other cats.’

- (28') *Dono punya seekor kucing putih dan seekor kucing hitam. Kucing putih
Dono has a cat white and a cat black cat white
ini lebih lama dipelihara Dono dibandingkan kucing-kucing lainnya.
ini more long taken care Dono compared cats other
'Dono has a white cat and a black cat. Dono has taken care of this white cat
for a longer time compared to his other cats.'

In sentences (28) and (28') there is an obvious contrast. The more recently mentioned cat is the black cat; *ini* can only be used to refer to this particular cat. When *ini* is combined with the earlier expression *white cat*, the sentence becomes ungrammatical (28'). With *itu*, however, the situation is very different. *Itu* can be used to denote either the black cat or the white cat.

- (28'') Dono punya seekor kucing putih dan seekor kucing hitam. Kucing hitam
Dono has a cat white and a cat black cat black
(kucing putih) itu lebih lama dipelihara Dono dibandingkan kucing-
cat white itu more long taken care Dono compared cats
kucing lainnya.
cat other
'Dono has a white cat and a black cat. Dono has taken care of the black
cat/white cat for a longer time compared to other cats.'

This data shows that there are definite contrasts between *ini* and *itu*, with *ini* behaving much more like a demonstrative, while *itu* patterns much more like a definite article. One of my language consultants also argues that the following sentence is ungrammatical or odd with *ini* because the fact that the bicycle has been stolen means that it is no longer there. Thus, one cannot refer to it by way of *ini* anymore. This suggests that *ini* is very strictly related to location, but *itu* is used in a totally different way.

- (29) Ali baru beli sepeda baru, tapi sepeda itu (*ini) sudah dicuri orang
Ali just buy bicycle new but bicycle DEF DEF already stolen person
kemarin.
yesterday
'Ali just bought a new bicycle, but that bicycle has been stolen by someone
yesterday.'

Interestingly, the same kind of contrast exists in English between *this* and *that*. *This*, for example is only felicitous in cases where there are contrastive differences.

- (30) I saw one quilt which was quite abstract, with lots of asymmetric diagonals. **Another** was more traditional, worked in an old Amish pattern. **This quilt** was less busy than the other, but just as bold.

(Roberts 2002)

English *this* is also not the preferred or default anaphoric article, as is the case in Indonesian.

- (31) I saw a movie about frogs. This movie about frogs was really good.⁴

Many scholars (Roberts (2002), Wolter (2006)) have also discussed in detail the use of demonstrative *that* as a special subtype of definite article. Roberts (2002) also observes, however, that in English *that*, with two equally possible antecedents, is unacceptable as an anaphoric definite description, while a demonstrative description necessarily refers to the most recently mentioned potential antecedent.

- (32) A woman_i entered from stage left. Another woman_j entered from stage right.
a. #The woman/That woman_j/She_j was carrying a basket of flowers.

Thus, we see that while *that* can be used as a definite article, there are instances where it shows more demonstrative properties (as seen in (32)). In Indonesian, however, the use of *itu* instead of *-nya* does not improve on the grammaticality of the sentence. Below I reproduce sentence (32) in Indonesian. This means that *itu* “fails” to function as a demonstrative that refers to the most recently mentioned potential antecedent.

- (32') Seorang wanita masuk dari kiri. Seorang wanita_j lain masuk dari kanan.
A woman enter from left a woman other enter from right
'A woman entered from the left. Another woman entered from the right.'
a. #Wanita-nya/ #Wanita itu membawa satu keranjang bunga.
Woman – nya woman itu bring a basket flower
'The woman / that woman brought in a basket of flower.'

Thus, it seems that *itu* is not really synonymous with the English demonstrative *that*, and behaves a little more like the English definite article *the*.

3. Num+Class, the Weak Article *-nya* and the Possessives

Next I reproduce my earlier example illustrating that co-occurrence of the weak article *-nya* and the Num+Class (which can also signal indefiniteness) is impossible.

⁴ This sentence was suggested to me by my committee member Molly Diesing.

- (33) *satu (buah) matahari -nya
one class. sun DEF
'the one sun'

Since *-nya* indicates absolute or situational uniqueness, this complementarity is not surprising. A DP cannot be both indefinite and unique in a specific situation, thus, the Num+Class complex that can also indicate indefiniteness cannot co-occur with the weak definite article *-nya*. Syntactically, this begets the question of whether or not there is a need to generate two separate maximal projections at all. Of course, this incomplementarity can simply be due to the semantics. However, even when we have numbers other than one occurring with the classifier, we still cannot have Num+Class co-occurring with the weak article *-nya*. Below I reproduce my earlier sentence from the introduction to illustrate this.

- (34) *Dua orang presiden -nya baru mendarat di Soekarno-Hatta.
two class. president DEF just landed at Soekarno-Hatta.
'The two presidents just landed at Soekarno-Hatta (the Jakarta airport).'
- (35) *Lima (buah) buku -nya mahal sekali
five class. book DEF expensive very
'The five books are very expensive.'

From this, it is clear that the co-occurrence impossibility we observe cannot be due to the semantics; and can only be syntactic. Assuming that the syntax avoids redundancy and favors simplicity, I propose that the maximal projection containing either the weak article *-nya* or the Num+Class complex is one and the same projection. Following my naming convention in the previous section, I will call this DP2.

Here, I reiterate that I treat the number and classifier in this language as a complex that occupies the Spec and Head position of the same projection. I have pointed earlier to the fact that the Num+Class complex forms the indefinite article as the main reason for believing that the number and classifier are contained within the same projection. However, the number and classifier are also the only elements that are to the left of the noun in the Indonesian DP.

- (36) Tiga buah bola merah besar saya yang bulat ini.
three (class) ball red big my that round this.
'These three red big ball that is round that belong to me'

Since they are the same projection (DP2) in the structure in my theory, they should branch in the same way. However, the DP word order is such that ClassP is prenominal and DP2 is post-nominal. Below I reproduce some example sentences from the introduction to illustrate this.

- (37) Bola merah-nya
ball red his/her
'His or her red ball / the red ball'
- (38) **Tiga (buah)** bola merah saya itu.
three class. ball red my that
'Those three balls of mine'

To account for the different branching of *-nya* and the Num+Class, I adopt Cinque's roll-up movement (see section 4 below).

Another co-occurrence impossibility that must be briefly addressed here is that of the weak definite article *-nya* and other possessives.

- (39) *Jam tangan-nya kamu.
Watch DEF you.
'The your watch'

The possessive can, however, co-occur with the Num+Classifier.

- (40) Sebutir telur Rina.
A egg Rina
'Rina's one egg.'

As such, I cannot posit that the possessive is also in the DP2 projection. There needs to be a separate DP projection for the Possessives. I propose that the weak article *-nya* that may itself function as a possessive is always base generated in this possessive projection. If *-nya* moves up into DP2, then it gets interpreted as the weak article. This is why we cannot get the possessive and the weak article co-occurring together.

4. Cinque's Roll Up Movement

Cinque (2005) utilizes Kayne's LCA in deriving the attested and unattested orders within the DP cross-linguistically. Cinque argues that elements that modify the noun are base generated pre-nominally in the strict order of Dem > Num > Adj > N. These maximal projections, in his theory, are also uniformly left-headed; DP-internal configurations other than the one base generated are derived by movement.

The following are the principles of Cinque's analysis:

- (a) The underlying hierarchical order in the extended projection of the noun is Agrw>W>AgrX>X>AgrY>Y>N where Y hosts AP in its specifier, X hosts NumP in its specifier and W hosts DemP in its specifier;
- (b) all (relevant) movements move a subtree containing N;
- (c) all movements target a c-commanding position;

- (d) All projections are modelled on the template (Kayne (1994)): [XP Spec [XP X0 Compl]]

To derive all and only the attested order of DP configurations, Cinque has to make a few more assumptions:

- (A) Merge order of the elements is Dem>Num>A>N
 (B) Parameters of movements:
 (i) No movement (unmarked)
 (ii) Movement of NP + pied-piping of the whose picture type (unmarked)
 (iii) Movement of NP without pied-piping (marked)
 (iv) Movement of NP + pied-piping of the picture of who type (more marked)
 (v) Total (unmarked) vs. partial (marked) movement of NP with or without pied-piping
 (vi) Neither head movement nor movement of phrase not containing NP is possible

Cinque's theory essentially involves different kinds of phrasal movements that move in succession through the Spec position of Agr in a roll-up fashion.

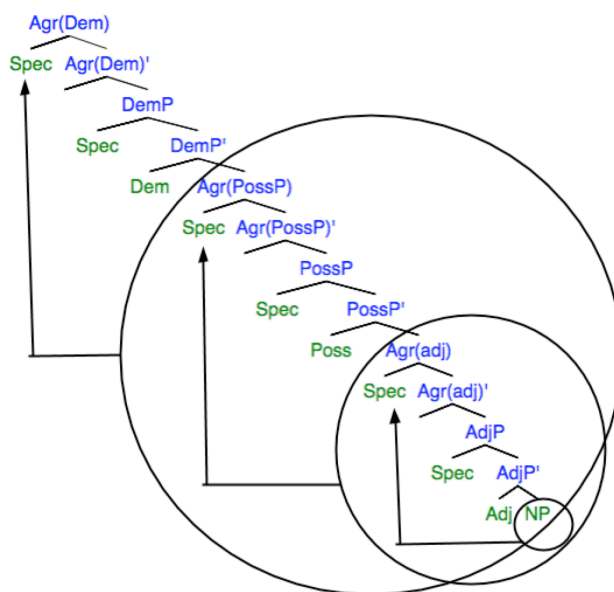


Figure 2: Cinque's roll-up movement

Under Cinque's theory, the fact that numbers and classifiers are prenominal while other elements are post-nominal means that Indonesian has to employ two types of pied-piping movements:

- (ii) Pied-piping of constituents of the whose picture type that is responsible for the postnominal order of elements that is the mirror image of the prenominal order seen in English.
 (iv) Pied-piping of constituents of the picture of who type that is responsible for the placement of numbers and classifiers prenominally.

If, as I mentioned earlier in the previous section, number and classifier are contained within the same projection (and Number is specifically in the Spec Agr of the classifier projection), then the Spec of Agr(ClassP) is always necessarily filled by the numeral; and we can arrive at a natural explanation of why movement type (ii) cannot happen; and why movement type (iv) must happen whenever Classifiers are present: because the presence of the numeral blocks movement type (ii) in Cinque's theory.

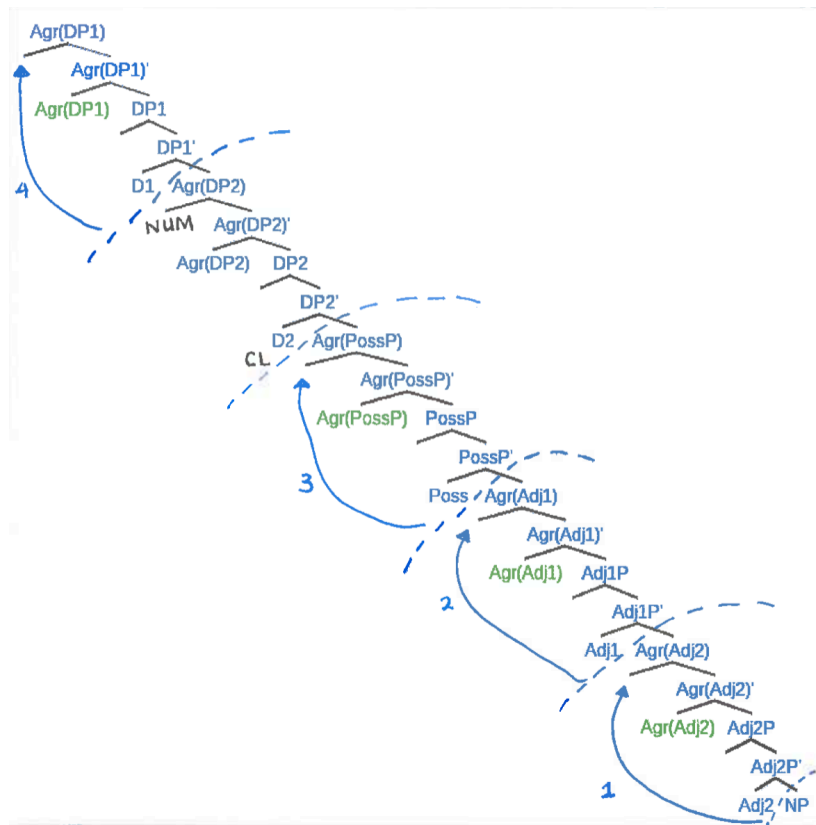


Figure 3: Roll-over movement for Indonesian with Num+Class

In the event that DP2 is filled with *-nya* instead of number and classifier, the Agr of DP2 is not filled with anything; and the roll-up movement can go through the Agr of DP2 without any hindrance, placing *-nya* to the right of the noun.

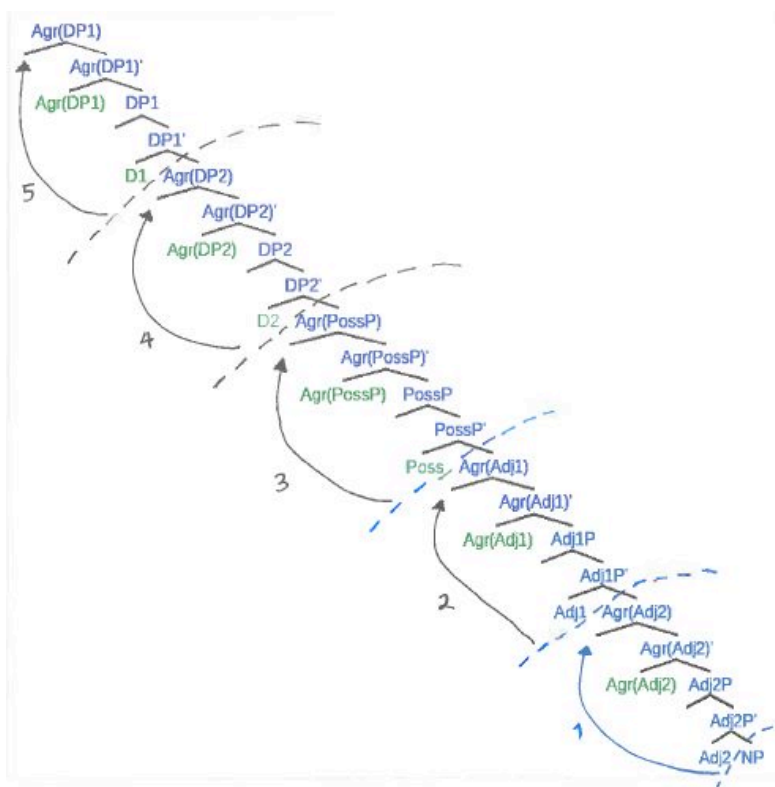


Figure 4: Roll-up movement for Indonesian DP with *-nya* in DP2.

We see thus that Cinque’s derivational account provides an explanation for the positions of the Num+Class and *-nya* in the syntax. A base generation account (Abels & Neeleman 2012) would, in contrast, find it difficult to explain the data.

5. Conclusion

Following the line of analysis adopted by other proponents of a layered categorical DP (Borer (2005), Heycock & Zamparelli (2003), et al), I propose 3 layers of DP in Indonesian: DP1, DP2; and the Possessive Phrase. I do not, however, make the claim that this structure should be universal. Instead, it is fully possible that this is a syntactic parameter, similar to what Boskovic (2008) and Despic (2011) have suggested for DP and DP-less languages.

Further research needs to be done, especially since I have omitted quantifiers from the discussion of Indonesian determiners in this paper. Also, while a preliminary comparison between A&N’s theory and Cinque’s theory seems to favor Cinque’s theory in general, I have not been able to provide solutions to the problem of Cinque’s heavy reliance on Agr projections and the assumptions of cartographic theory.

References

- Abels, Klaus and Ad Neeleman. 2009. Universal 20 without the LCA. *Merging features: Computation, interpretation and acquisition*, eds. Josep M Brucart, Anna Gavarró, and Jaume Solà, 60–79. Oxford University Press.
- Abney, Steven Paul. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT, Cambridge, Ma.
- Borer, Hagit. 2005. Structuring sense, vol. 1: In name only. Oxford University Press
- Bošković, Željko. 2008. What will you have, DP or NP? *Proceedings of the Northeast Linguistic Society 37*, eds. E. Elfner and M. Walkow, 101-114. Amherst, Ma.
- Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36:315–332.
- Despić, Miloje. 2011. Syntax in the absence of determiner phrase. Doctoral dissertation. University of Connecticut.
- Heycock, Caroline & Zamparelli, Roberto. 2005. Friends and colleagues: Coordination, plurality and the structure of DP. *Natural Language Semantics* 13:201-270
- Ihsane, Tabea & Puscas, Genoveva. 2001. Specific is not Definite. *Generative Grammar in Geneva* 2:39-54
- Roberts, Craige. 2002. Demonstratives as definites. *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*, eds. K. van Deemter and R. Kibble, 89–106. Stanford, Ca.: CLSI Press.
- Schwarz, Florian. 2009. Two Types of Definites in Natural Language. Doctoral Dissertation, U of Massachusetts Amherst.