JSEALS Special Publication No. 5

PAPERS FROM THE AUSTRONESIAN
FORMAL LINGUISTICS ASSOCIATION 25

AFLA
25

Edited by
Henry Y. Chang
Hui-chuan J. Huang
The Journal of the Southeast Asian Linguistics Society publishes articles on a wide range of linguistic topics of the languages and language families of Southeast Asia and surrounding areas. JSEALS has been hosted by the UH Press since the beginning of 2017.
Contents

Introduction to Special Issue: Papers from AFLA 25 ...................................................... iv

From the JSEALS Editor-in-Chief ........................................................................................ vi

Ways of talking about the past: The semantics of –in- and \=in in Bunun ......................... 1
Sihwei Chen and Haowen Jiang

More on Kanakanavu word-level prosody: Cyclic and postcyclic processes ............... 22
Yi-Yang Cheng

The distribution of the Māori genitive relative construction ..................................... 47
Michael Socolof and Junko Shimoyama

A presentational construction in Indonesian ................................................................. 60
Carly J. Sommerlot

Untangling the Tagalog clitic cluster ............................................................................. 80
Russell Tanenbaum

Two types of noun incorporation in Northern Paiwan .................................................. 97
Chunming Wu

Malay verbal reduplication with the məN- prefix ......................................................... 118
Meng Yang and Deborah J.M. Wong
INTRODUCTION FROM THE VOLUME EDITORS

This volume grew out of the 25th annual meeting of AFLA held at Academia Sinica, Taiwan, 10-12 May, 2018. This occasion featured 24 talks on the generative study of Austronesian linguistics. Seven of the talks were written up into full papers and submitted to us for consideration of publication. Each of them was reviewed by two external anonymous reviewers, revised in accordance with the reviewers’ comments, and finally included here.

Contributions in this volume cover a wide variety of topics in Austronesian linguistics. Chen and Jiang argue that in Bunun, -in- is an existential past tense marker while =in is a change-of-state marker at the discourse level, in contrast to the dominant view in the literature. Focusing on the prosody of Kanakanavu, Cheng spells out a number of phonological conditions and identifies the morphemes that could either attract or repel prominence. Socolof and Shimoyama propose a split-ergative analysis of Māori genitive relative construction while showing that this construction is more widely distributed than generally described. Sommerlot’s article shows that the ber-V-nya constructions in Indonesian do not fit into any functions of these affixes in previous descriptions and they instead resemble a type of presentational-there construction. Tanenbaum adopts a syntactically-grounded account of Tagalog second-position clitics, based on obligatory V-to-C head movement. Wu explores the constructions of noun incorporation (NI) in Northern Paiwan, including both lexical and syntactic NI, and examines their morphosyntactic behaviors. Yang and Wong study how Malay məN- prefixation interacts with reduplication and propose a new markedness constraint against word-initial nasals to account for the data.

We thank the JSEALS Editor-in-Chief, Mark Alves, for his unwavering support. We are also grateful to the anonymous reviewers for their insightful comments that have led to significant improvements of the quality of the articles. Liok-san Ng’s editorial assistance is also acknowledged. Hopefully, this volume can contribute to a better understanding of Austronesian languages and the advancement of Austronesian linguistics.

Henry Y. Chang
Taipei, Taiwan

Hui-chuan J. Huang
Taipei, Taiwan

22 December 2019

Editors
Henry Y. Chang (PhD) is a Research Fellow in the Institute of Linguistics at Academia Sinica in Taiwan. He received his PhD from National Tsing Hua University, Taiwan in 1997.

Hui-Chuan J. Huang (PhD) an Associate Research Fellow in the Institute of Linguistics at Academia Sinica in Taiwan. She received her PhD from University of Minnesota, USA in 1997.
FROM THE JSEAL EDITOR-IN-CHIEF

This is the fifth JSEALS special publication. The goal of JSEALS special publications is to share collections of linguistics articles, such as select papers from conferences or other special research agendas, as well as to offer a way for linguistic researchers in the greater Southeast Asian region to publish monograph-length works.

This volume contains seven papers from the AFLA 25 conference. The languages covered in this volume are spoken in Taiwan, Indonesia, and Oceania. This vast extension on “Greater Southeast Asia” is the case because, though Austronesian has an insular Southeast Asian presence and origins, it has famously extended quite far through Austronesian seafaring skill, making it necessary to include languages in this tremendous geographic range. The papers vary in the topics, including phonology, morphology, and syntax, making this a solid contribution to theoretical linguistics in general.

We are very pleased that JSEALS is able to contribute to the sharing of quality linguistic research in both mainland and insular Southeast Asia.

Mark J. Alves
Rockville, Maryland
January 15, 2019
WAYS OF TALKING ABOUT THE PAST:

THE SEMANTICS OF -IN- AND =IN IN BUNUN

Sihwei Chen
Academia Sinica
sihweichen@gate.sinica.edu.tw

Haowen Jiang
Peking University
jhw@pku.edu.cn

Abstract
This paper assesses the semantics of two homophonous markers, -in- and =in in Isbukun Bunun (Formosan, Austronesian). While both markers are used to express some kind of anteriority reading, there is no consensus about their semantic category. By presenting new empirical evidence, we offer an alternative analysis for both markers. We argue in detail that -in- is an existential past tense marker, which is in line with Jeng (1999) and De Busser (2009) but differs in many important details. We also suggest that =in is a discourse-level change-of-state marker, in contrast to the dominant view. This study not only suggests that temporal components can be associated with different categories cross-linguistically, but also has implications for analyzing similar markers in other Formosan languages.

Keywords: past tense, perfect aspect, inchoativity, Isbukun Bunun, existential past

ISO 639-3 codes: bnn

1 Introduction
This paper assesses the semantics of two homophonous markers, -in- and =in in Isbukun Bunun (Formosan, Austronesian). The former is an infix and has an allomorphemic variant -i- while the latter is an invariant enclitic attachable to various lexical words. We choose to look at the two markers for a number of reasons. First, their phonological identity raises the question of how their use differs. Second, while both markers are used to express some kind of anteriority reading, there is no consensus about their semantic category. The infix -in- has been analyzed or described differently as marking past tense (Zeng 1986; Huang 1997; Jeng 1999; De Busser 2009), experiential (perfect) aspect (Lin 1997; Huang & Shi 2016), perfective aspect (Zeitoun et al. 1996), or telicity (Lin 2010). On the other hand, the enclitic =in is predominately described as a perfect aspect marker (Zeitoun et al. 1996; Huang 1997; Lin 1997; Jeng 1999; Lin 2010), but yet it has also been considered to be a perfective (De Busser 2009). Finally, investigating the semantics of -in- and =in in Bunun has implications for studying similar phenomena in other Formosan languages, including widespread cognates of -in- and diverse forms of enclitics loaded with a function comparable to that of =in.

By presenting new empirical evidence, we offer an alternative analysis for both markers. We argue that -in- is an existential past tense marker, which is in line with Jeng (1999) and De Busser (2009) but differs in many important details, and =in is a discourse-level change-of-state (COS) marker, in contrast to the dominant view. The two markers differ in how they express anteriority: while the past tense -in- lexically specifies that there is a past time at which the described event holds, the anteriority effect with =in arises due to a pragmatically conditioned change of state. Our findings also show that -in- and =in possess decomposable semantic features common to temporal operators in other languages, thus bearing implications for cross-linguistic studies on the semantics of tense/aspect.

1 We would like to thank our language consultant Hanaivaz Takistaulan (born in 1951) for her grasp of and devotion to the Bunun language, two anonymous reviewers for their encouraging and valuable comments, and finally the audience at AFLA-25 for all the issues brought to our attention.

2 When the first syllable of the verb root contains a consonant and the vowel /a/, -i- instead of -in- is inserted after the vowel /a/ (Zeng 1986). For example, after undergoing in-infiction, ma-sabah ‘AF-sleep’ becomes ma-<i>sabah ‘AF-<E.PST>sleep’.

Copyright vested in the author; Creative Commons Attribution License
The organization of this paper is as follows. In the rest of this section, we introduce the Bunun language and the dialect which our data is based on (Section 1.1), review previous descriptions of these two markers in the literature (Section 1.2), and elucidate our methodology of elicitation (Section 1.3). Next, Sections 2 and 3 respectively discuss the empirical evidence for our analysis of the infix \(-\text{in}\) and of the enclitic \(=\text{in}\). Section 4 summarizes this paper and explores the implications of our findings.

1.1 Language background

Five dialects and three branches of Bunun are generally recognized (Li 1988): Northern (Takituduh & Takibakha), Central (Takbanuaz & Takivatan), and Southern (Isbukun), which historically correspond to five family clans. The heartland of the Bunun language is believed to be central Taiwan, specifically mountainous regions in Nantou County, from which clan members migrate to other parts of Taiwan, mostly Hualien and Taitung County in the east and Kaohsiung City in the south. The Southern dialect is the most widely distributed one, followed by Central and then Northern dialects. Possibly due to this geographical distribution, the Southern dialect is also the most researched one, again followed by Central and then Northern dialects. Based on speakers’ self-report, Southern speakers are better understood by Northern and Central speakers than the other way around. In this study, we specifically deal with Isbukun Bunun spoken in Namasia District of Kaohsiung City. Hence, all the Bunun data presented in this paper comes from Isbukun unless otherwise indicated. All cited data are modified for the sake of orthographic and glossing consistency.

1.2 Previous descriptions of \(-\text{in}\) and \(=\text{in}\)

Different terms have been proposed for the infix \(-\text{in}\) and enclitic \(=\text{in}\) based on different Bunun dialects. In Zeitoun et al.’s (1996) typological investigation of Formosan temporal/aspectual systems, reflexes of the Proto-Austronesian infix \(*-\text{in}\) in several Formosan languages, including Isbukun Bunun, are claimed to be perfective aspect markers. Their main reason is that sentences with \(-\text{in}\) “must be interpreted as referring to a past and completed situation” (p.29), and the event is “viewed in disconnection with S[peech] T[ime]” (p.45). The same view is adopted by many subsequent studies (e.g. Su 2008; Li 2010; Hsieh 2011; Jiang 2012). However, the perfective proposal is probably not on the right track, since, as pointed out by De Busser (2009:234), sentences marked with \(-\text{in}\) need not be interpreted as completed (see (3) below). Moreover, the disconnection with speech time is what is commonly found with a past tense rather than with a perfective aspect, as in English and many other languages (e.g. Musan 1997; Lin 2007; Altschuler & Schwarzschild 2013; Thomas 2014; Bochnak 2016; Cable 2017). In Section 2.1.1 we shall present evidence against a perfective analysis of \(-\text{in}\).

As for the enclitic \(=\text{in}\) in Bunun, Zeitoun et al. treat it as a perfect aspect marker, on the grounds that “with \(-\text{in}\), the event is on-going and has a certain relevance (resultant state/perfect) at speech time” (p.45). This claim is illustrated by (1), the free translation of which suggests that the drinking event continues from a past time up to the utterance time.

\[
\begin{align*}
\text{(1)} & \quad \text{Hud} = \text{in} & \text{saikin} & \text{danum}. & \quad^4 \\
\text{drink} = \text{IN} & \quad 1\text{SG.NOM} & \text{water}
\end{align*}
\]

‘I have been drinking water.’ (Zeitoun et al. 1996:45)

However, as we will show in Sections 3.2 and 3.3, \(=\text{in}\) always induces a change of state unless it is used to convey a contrary-to-expectation reading. In other words, in contexts without such a reading, sentences with \(=\text{in}\) cannot express a continuous/universal-perfect reading (compare (1) to (41) below). Based on their

3 Zeitoun et al. (1996) also claim that perfectivity in other Formosan languages is denoted by various enclitics rather than \(-\text{in}\) (e.g. \(=\text{la} [\text{la}]\) in Nanwang Puyuma, \(=\text{tu} [\text{tu}]\) in Amis, and \(=\text{nga} [\text{ŋa}]\) in Budai Rukai). However, our study of the Bunun \(=\text{in}\) suggests a re-examination of the perfective account in these languages (see Sections 3.2 & 4.2).

4 We adopt the conventional orthography and follow the Leipzig Glossing rules. Additional gloss abbreviations are as follows: AF ‘Actor Focus’; CONJ ‘conjunction’; COS ‘change of state’; E.PST ‘existentia past’; EPIS ‘epistemic’; HUM ‘human’; INTJ ‘interjection’; LF ‘Locative Focus’; LNK ‘linker’; NEUT ‘neutral case’; NPST ‘non-past’; PF ‘Patient Focus’; PRT ‘partitive’, RF ‘Referential Focus’.

5 We do not discuss “resultant state/perfect” mentioned by Zeitoun et al. since it is unclear what it refers to and no illustrative examples are provided.
proposal for -in- and =in, Zeitoun et al. consider Bunun to be the only Formosan language that grammatically distinguishes perfective aspect and perfect aspect (marked by the infix -in- and the enclitic =in, respectively). However, these claims are made without sufficient and conclusive evidence; in this paper, we provide new evidence for alternative proposals.

In his dedicated study on tense and aspect in Bunun, which covers one dialect from each of the three branches (i.e., Takituduh, Takbanuaz, and Isbukun), Jeng (1999) argues that the infix -in- is a marker for past tense. His main evidence comes from the fact that the marker is compatible with past adverbs, but not with present or future ones, as shown in (2).6

(2)  M<in>is’av naia {sangan / habas / *laupaku / *sanganin}.
    <IN>AF.drink.wine.at.a.festival 3PL.NOM a.while.ago / in.the.past / now / later.on
    ‘They drank wine at a festival {a while ago/in the past/*now/*later on}.’ (Jeng 1999:460)

However, according to Jeng, sentences marked with -in- can alternate with those without -in- if the context clearly indicates a past reference time (p.462). There is, however, strong evidence that the two forms are used in different contexts, one as an existential tense and the other as a pronominal one; this will be presented in Section 2.3. Jeng also argues against treating -in- as a perfective aspect, since -in- can co-occur with aspects that convey a progressive or durative reading (see e.g., (9) and (10) below); yet he states that -in- is a past tense that implies a default perfective meaning (i.e., a situation completed as a whole in the past) (p.473). As discussed above, the use of -in- is not always associated with completion of events (De Busser 2009), and hence -in- cannot encode perfectivity.

Like Zeitoun et al., Jeng (1999) claims that the enclitic =in is a perfect aspect, but on different grounds. While Zeitoun et al. define perfect aspect by continuation of an event to the speech time, Jeng’s proposal is based on anteriority of an event. As he puts it, =in “indicates the inception of a situation (a state or action) prior to a reference time and the situation may be either completed or continue to be in existence or in progress relevant to the reference time” (p.475; emphasis ours). Jeng further considers the co-occurrence of =in and the future marker na= to be interpreted as the beginning of an event prior to a future reference time (e.g., ‘They will have begun to eat rice’, similar to the use of English future perfect). However, we will show that contrary to this claim, no anteriority reading arises with a future reference time (see Section 3.1).

De Busser (2009) argues, based on Takivatan data, that the infix -in- is always associated with past tense semantics (with resultative undertones, see e.g., (11) below).7 A piece of data he provides, which constitutes crucial evidence against a perfective analysis, is that the event marked by the infix -in- need not be completed, as in (3).

(3)  We were not yet at that location, because…
    T<in>as’i kaku tudip’az Sipun=ang.
    <IN>build school that.time-NMLZ Japan=still
    ‘The school was still being built when the Japanese were still there.’ (Takivatan Bunun, De Busser 2009:234)

We will show in Section 2.3 that, in line with both Jeng (1999) and De Busser (2009), a past tense analysis better captures the properties of -in-, for which we further provide a formal account.

As for the enclitic =in, De Busser differs from most other researchers by analyzing it as a marker for perfective aspect for the reason that it indicates completion of events. However, De Busser’s definition of completion means that “an event has finished or that a certain endpoint has been reached” (p.224); in this sense, it not only refers to the culmination of telic events but also a termination point of any events. For example, (4)

---

6 On a related note, Li (2018:289) finds it hard to reconcile the fact that the alleged perfective Isbukun -in- “is incompatible with an event or a situation that does not take place in the past” and that perfective markers in languages like Palauan can combine with present-tense markers.

7 Both Jeng (1999) and De Busser (2009) propose that -in- as a past tense is not always deictic but can be relative to the time of a past matrix clause (or to a past narrative time). Since our data are mostly centered on unembedded clauses, we leave the issue of relative tense for future research.
is intended to show that the washing event terminates (as suggested by the translation—part of my body was washed).

(4)  
Laupang='ak  taldanav=in.
now=1SG.NOM  AF.wash=IN
‘I just finished washing (part of my body).’ (Takivatan Bunun, De Busser 2009:225)

Moreover, De Busser claims that =in “certainly never indicates continuation” (p.229). These generalizations are mostly based on translation, and because there are no clear entailment tests, it is unclear how one can be sure whether culmination or termination is involved, or whether events continue or not. As we shall see in Section 3.2, once all predicate types are taken into consideration, a clear picture emerges: =in is used for not only culminated events, but also inceptive ones, depending on the type of event; we will also demonstrate that event inception can be associated with ongoing or forthcoming events.8

Lin (2010) adopts a perfect analysis of the enclitic =in, but argues that the infix -in- is a telicity marker, turning atelic events into telic ones. According to Lin, types of object that verbs take determine the telicity of the predicates in Bunun: they are telic with a numeral modifier, and atelic otherwise. Lin shows that when marked by -in-, an atelic predicate cannot be continued with a sentence which states that the event continues up to the utterance time, as shown in (5).

(5)  
Ma-<i>kulut  is  lukic  ca  banaz  takna’,  #  unto  ma-kulut=ang  cia  ca
AF-<IN>cut  OBL  tree  NOM  man  yesterday  then  AF-cut=still  3SG.OBL NOM
man  now
Intended for ‘The man chopped a tree yesterday, and the man is still cutting it now.’
(Takituduh Bunun, Lin 2010:77)

However, this example only shows that the chopping event has terminated. Given that termination is typologically possible for both perfective telic and atelic events (see e.g., Singh 1998; Altshuler 2014), this test cannot be taken as an entailment test for telicity/culmination of telic events. What typically diagnoses whether telic events reach their endpoints (or that atelic events are shifted into telic ones) is the possibility of coordinating with a statement that the event is not completed; this is nevertheless not shown in Lin’s discussion. We leave a detailed investigation of lexical aspect in Bunun to another occasion, and simply note that a telicity analysis does not straightforwardly capture the properties of -in- to be discussed in this paper.

Finally, in their recent reference grammar of Isbukun Bunun, Huang and Shi (2016:117-119) consider the infix -in- to be a subset of perfect exclusively used for experiential readings, and equate it with the Mandarin -guo, a marker similarly used for experiential readings (see e.g., Lin 2007; Wu 2008). Nevertheless, not all of their examples can be interpreted as experiential or translated with -guo. For example, the sentence in (6), in which -in- modifies a stative verb, can only be translated with a past tense reading, and the original Mandarin translation is incompatible with -guo. We will argue below that the experiential reading is simply an instantiation of existential past tense without domain restriction (Sections 2.1.2.2 and 2.4).

(6)  
'<in>i-lumah   sangan  a  tama=a.
<IN>AF.be.at-home a.while.ago NOM father=DIST.NOM
‘Father was home a while ago.’ (Original translation in Mandarin, Huang and Shi 2016:138)

8 In addition to completion, De Busser (2009) identifies three other functions of -in-: resultative meaning, change of state, and anteriority. Without offering much evidence, he assumes that completion and resultative meaning are “instantiations” of perfective and change of state and anteriority are only its “functional extensions” (p.230). Note that in De Busser’s discussion the resultative meaning does not mean the result state of a telic event but “an event as being the result of some other event” (p.225), and the anteriority of =in means that the event of a temporal clause is anterior to that of the main clause.
As for the enclitic =in, Huang and Shi term it perfect but analogize it to the Mandarin sentence-final particle le, which has been given analyses beyond a temporal category (e.g., Soh 2009). They explicitly state that the reading that =in yields is determined by the predicate it modifies: inceptive readings with activity predicates, completive readings with achievements, and change-of-state readings with statives. Our data not only support this more fine-grained characterization but also complement it with accomplishment predicates; these will be discussed in Section 3.2.

This section has summarized the major proposals for the two Bunun markers over the past two decades. The infix -in- shares features with past tense—its use is restricted to the past and reveals disconnection effects from the present time. It has been considered to be a perfective or an experiential perfect aspect, but the decision is often exclusively based on translation without independent semantic tests. By contrast, the enclitic =in has a rather different property: it refers to an anterior event, and relates it to the current situation; it also shows interactions with different types of predicate. Against the backdrop of these studies, we will argue in this paper that neither of these markers denotes the semantics of perfective or perfect.

A final remark is on the interaction of -in- and the four ways of marking focus (AF m-/ma-; PF -un; LF -an; RF is-/s-). Blust (1998:347) notes that “while PAn *-in- and its reflex in many daughter languages may co-occur with the IF [RF in this study], AF, or LF affixes, it has a zero allomorph with the PF suffix.” In other words, reflexes of PAn *-in- in many languages have a portmanteau function for both tense/aspect and PF marking. Interestingly, we have identified in Jeng and Ispalidav’s (2016) Isbukun Bunun Dictionary word forms and/or example sentences of -in- co-occurring with all four focus affixes, including the PF. We illustrate this distribution with unmarked word forms followed by their in-marked counterparts: AF manah vs. m<in>anah ‘to shoot, hunt’; PF ludah-un vs. l<in>udah-un ‘to hit, strike’; LF pa’anak-an vs. pa<i>-anak-an ‘to beat up’; RF is-pinang vs. s<in>-pinang ‘to sow (seeds)’. At this point, it is unclear whether the PF form <in>V-un in Bunun is sporadic or entrenched.

1.3 Methodology
Unlike previous studies, most of our data are embedded under a specific context and accompanied with grammaticality and felicity judgments as well as the consultant’s comments (see Matthewson 2004 for the procedures for contextual elicitation). Moreover, these data are augmented by a storyboard based on ‘Miss Smith’s Bad Day’ (Matthewson 2014). The procedure of eliciting the storyboard is as follows: we first illustrated the story frame by frame to the speaker in Mandarin, and then we recorded the speaker’s retelling of the story. In subsequent sessions, we then went through the recorded story with the speaker to transcribe the story and to conduct follow-up elicitation. For the details and benefits of utilizing storyboard in elicitation, the reader is referred to Burton and Matthewson (2015).

2 The semantics of -in-
We argue that the infix -in- is not any kind of aspect but a past tense marker. The examination of -in- against the properties of several aspectual categories shows that an aspectual analysis is unattainable (Section 2.1 and 2.2). Instead, the properties of -in- are all explainable if -in- is analyzed as a past tense and, more precisely, as an existential past tense, following the diagnostics proposed in Chen et al. (2017) for Javanese and Atayal (Section 2.3). Section 2.4 provides a formal analysis.

---

9 Huang and Shi (2016) also consider that the suffix -nahtung marks perfect just like the enclitic =in, but they do not show what types of perfect semantics the two markers each encode.

10 While translation can be useful at different stages of elicitation, it is not a sufficient method of exploring meaning (cf. Matthewson 2004).


12 Blust (1998) points out that Thao is unique among Formosan languages since it allows the combination of -in- and PF. We are not sure how robust the same combination is in Bunun, but for our purpose the crucial point here is that Bunun -in- is compatible with various focus markers.

13 The original storyboard can be found at http://totemfieldstoryboards.org/stories/miss_smith/.
2.1 Not an aspect marker

In this section, we present evidence that the infix -in- cannot be analyzed as a perfective aspect or a(n) experiential (perfect) aspect.

2.1.1 Not a perfective aspect marker

Jeng (1999) argues that -in- is not a perfective aspect but a past tense based on the past temporal restriction of -in- and the compatibility of -in- with viewpoint aspects including an imperfective one. Our data confirm these claims. In what follows, we first illustrate these two pieces of evidence with our own data and then provide further support for a past tense analysis.

First, the occurrence of -in- is restricted to a past reference time. As shown in (7), the presence of -in- in the sentence (realized as -i- here, see fn. 2) is inconsistent with the presence of a present-time adverb (see also (2) above).

(7) \(Ma<\text{E.PST}>\text{sabah} \text{ saia } \{\text{takna } / \ast \text{laupakadau}\} \text{ sia } \text{ sapalan}=\text{cia}.\)

\(<\text{E.PST}>\text{AF.sleep} \text{ 3SG.NOM yesterday } / \text{now} \text{ LOC bed}=\text{DIST.OBL}\)

‘He slept on that bed yesterday.’ / ≠ ‘He is sleeping on that bed now.’

Future-time reference in Bunun requires the proclitic \(na=\), as in (8)a,\(^{14}\) and the presence of -in- in such sentences is inelicitous, as in (8)b (see also Huang 1997:380).\(^{15}\)

(8) Context: “Can I come over at 3 pm. tomorrow?” “No, 3pm. is not good because …”

a. *(Na)=ma-pa-tal’isuh saikin Aping=cia.

\(\text{FUT}=\text{AF-CAUS-bathe} \text{ 1SG.NOM Aping}=\text{DIST.OBL}\)

‘I will be giving Aping a bath.’

b. \(Na=ma-\ast<\text{E.PST}>\text{pa-tal’isuh} \text{ saikin Aping=cia}.\)

\(\text{FUT}=\text{AF}<\text{E.PST}>\text{CAUS-bathe} \text{ 1SG.NOM Aping}=\text{DIST.OBL}\)

‘I will be giving Aping a bath.’

Data such as these clearly show that -in- is temporally restricted and cannot be an aspect marker in the theory of Reichenbach (1947) and Klein (1994): such aspectual markers should be able to combine with a past, present, or future tense to yield a different reference time. The observed restriction to the past thus favors the proposal that -in- is a past tense rather than an aspect such as perfective. Nevertheless, perfectives in many languages (including the English simple past) can be pragmatically restricted to the past (Bennett and Partee 1978; Dahl 1985; Kamp and Reyle 1993; Giorgi and Pianesi 1997; Smith 1997). There is also a possibility that -in- is a past perfective marker. These possibilities, however, are ruled out by the fact that -in- can co-occur with Ca-reduplication (which has been considered to be an imperfective aspect, cf. Jeng 1999; Huang and Shi 2016), yielding a past habitual or past progressive reading:\(^{16}\)

---

\(^{14}\) In our data, \(na=\) is obligatory for the future. This fact, however, is inconsistent with the note made by De Busser (2009:212) based on Takivatan Bunun spoken in Hualien; that is, “[i]t is perfectly possible to have a future event without explicit irrealis marking [i.e., \(na=\)].”

\(^{15}\) The only case where -in- and \(na=\) can co-occur is when \(na=\) is interpreted as epistemic. For example, -in- in (i) marks a past scolding event that the speaker conjectures (see also Huang and Shi 2016:128) (but see Lin 2010:109; De Busser 2009:241).

(i) \(Na=h<\text{E.PST}>\text{aungun-an saia takna mas isaicia tu cina, aupa mahansu dahis}=a.\)

\(\text{EPIS}=<\text{E.PST}>\text{scold-LF} \text{ 3SG.NOM yesterday OBL 3SG.POSS LNK mother because AF.stink face}=\text{DIST.NOM}\)

‘He might have been scolded by his mother because he looks upset.’

\(^{16}\) Example (10) falsifies Huang and Shi’s (2016:125) claim that the combination of -in- and Ca-reduplication cannot be interpreted as progressive. We leave the difference between progressive readings with Ca-reduplication and those without for another occasion; this hinges on a detailed investigation of Ca-reduplication, which is beyond the scope of this paper.
(9) Context: Describing my grandma’s life before.
Saia masa makuang=ang lutbu hai, kaa ma<l>t<s>a->sabah.
3SG.NOM when.PST AF.bad=still body TOP just <E.PST><IPFV~>AF.lie.down
‘When she was still ill, she often just lay (on something).’

(10) Context: Describing what I was doing yesterday.
Ma-<E.PST>IPFV~>da-damu saikin haludun.
AF-<E.PST>IPFV~>catch 1SG.NOM cricket
‘I was catching crickets.’

These data would not be possible if -in- were analyzed as a perfective aspect or as encoding a perfective component.17

Another property which -in- shares with past tense rather than with perfective aspect is that when marking stative predicates, it induces an inference that the described state ceases to hold in the present.18 This is illustrated by (11).

(11) M<in>asmuh a saia habas.
<E.PST>AF.fat NOM 3SG.NOM before
‘(S)he was fat before (and is not fat now).’ (Huang & Shi 2016:120)

This property is similarly found with the past tense in English (e.g., Musan 1997; Althshuler and Schwarzschild 2013) and in other languages (e.g., Thomas 2014; Bochnak 2016; Cable 2017). By contrast, perfective statives are not typically interpreted in the past; instead, they often induce coerced readings with states, e.g., inchoative readings, at the present time (Bybee et al. 1994; Tonhauser 2006). If -in- were a perfective marker, stative predicates marked by -in- would be either impossible or result in a present inchoative reading.

Overall, the properties discussed above are all expected if -in- is analyzed as a past tense: Firstly, as a past tense, -in- exhibits a temporal restriction to the past. Secondly, -in- can co-occur with Ca-reduplication to yield past habitual or progressive readings. Lastly, stative predicates marked by -in- are interpreted in the past and share the cessation effect with past tenses in other languages.

2.1.2 Not an experiential (perfect) aspect
The infix -in- has a dominant experiential reading, as exemplified in (12), and for this reason it has been considered to be an experiential (perfect) aspect (e.g., Lin 1997, Huang and Shi 2016).

(12) M<in>uhalhal saikin sia lukis.
<E.PST>AF.fall 1SG.NOM LOC tree
‘I have once fallen from a tree.’ (Original translation in Mandarin, Huang and Shi 2016:57)

In this section, we first show that an analysis that equates -in- with the English perfect aspect will not capture the fact that -in- does not have any other uses of the perfect. Furthermore, -in- does not always give rise to an experiential reading, which calls into question the analysis of -in- as an aspect that exclusively marks experiential readings.19

---

17 Another type of evidence would be like the Takivatan Bunun example in (3) above, where sentences with -in- can express an ongoing reading typical of imperfective/progressive aspect. Data in Isbukun Bunun requires further work.
18 The cessation effect is described as a “past/present contrast” in De Busser (2009), according to which, sentences with -in- “express the idea that some past event is meaningfully different from the present situation” (p.235).
19 The infix -in- is sometimes described as an “experiential marker”, and a reviewer asks why we only examine a perfect analysis. However, the term “experiential marker” is usually given without an explanation or analysis, and it is not our purpose to speculate what analysis it refers to. What we argue here is that the experiential reading of -in- is only one of its many readings, all of which can be unified by existential quantification.
2.1.2.1 Nothing resembles the English perfect except for experiential readings

The infix -in- shares nothing in common with the English present perfect aspect except for an experiential reading. Table 1 lists eight properties of the English present perfect, each of which is to be examined against the Bunun -in-.

<table>
<thead>
<tr>
<th>Properties</th>
<th>-in-</th>
<th>have + p.p.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiential perfect</td>
<td>y</td>
<td>y</td>
</tr>
<tr>
<td>Adverbial restrictions</td>
<td>n</td>
<td>y</td>
</tr>
<tr>
<td>Current relevance</td>
<td>n</td>
<td>y</td>
</tr>
<tr>
<td>Lifetime effects</td>
<td>n</td>
<td>y</td>
</tr>
<tr>
<td>Recent past/hot news</td>
<td>n</td>
<td>y</td>
</tr>
<tr>
<td>Result state</td>
<td>n</td>
<td>y</td>
</tr>
<tr>
<td>Universal perfect</td>
<td>n</td>
<td>y</td>
</tr>
<tr>
<td>Anteriority (ET &lt; RT)</td>
<td>n</td>
<td>y</td>
</tr>
</tbody>
</table>

A brief note about this comparison is in order. We do not presuppose that the English present perfect is prototypical or standard; rather, the well-studied properties of the English perfect serve to discover the behaviors of the Bunun -in-. A similar method has been applied to many other languages (e.g., Mathewson et al. 2015; Bowler and Ozkan 2017; Chen 2017; Bertrand et al. 2017; a.o.), the result of which does not lead to an English-type perfect but reveals interesting cross-linguistic variation.

First, the experiential perfect reading of the English present perfect is similarly attested with the Bunun -in- but this is the only similarity between the two markers; they share none of the remaining properties in Table 1. In addition to (12) above, (13) is another example that illustrates the experiential reading. It is taken from our storyboard, where both Miss Savi’s question and Biung’s reply concern an event that happens at some point in one’s life.

(13) a. Miss Savi: Sima kamu sai-sian ludun mu-da~daan?
       Who 2PL.NOM <E.PST>go-DEM mountain AF.move-IPFV~road
       ‘Who has climbed (lit. gone and walked on) a mountain?’

               1SG.NEUT 1SG.NEUT AF.know 1SG.NOM go< E.PST>~DEM 1SG.NOM mountain
               ‘Me! Me! I know. I have climbed (lit. gone to) a mountain.’

In English, the present perfect cannot co-occur with adverbs that express a definite past time, as in (14)a, and this contrasts with the past perfect or a tenseless perfect in (14)b-c. However, in Bunun, the marker -in- is compatible with different types of past-time adverbial, for example, takna ‘yesterday’ in (15), tangusan tu buan ‘last month’ in (16), sangan ‘a while ago’ in (2) and (6) above, and habas ‘in the past’ in (2) above.

(14) a. *Chris has left York yesterday.          (Klein 1992:525)

       b. Uli had left on Tuesday.          (Katz 2003:147)

       c. Peter believes Steven to have gone to Boston last summer.  (Katz 2003:147)

20 The motion verb sai-sian is a suppletive form of ku-sian ‘go-DEM’ marked with -in-.
21 This adverbial restriction is absent with perfects in many other languages, even those closely related to English (Giorgi and Pianesi 1997). While we cannot conclude from this diagnostic, the full comparison does not point to a perfect analysis.
Events expressed by the English present perfect are known to bear certain relevance to the current context. This can be illustrated by one of Portner’s (2003) examples: the second perfect sentence in (17) is odd in the context where the illness of Mary has no significant result, but is acceptable when some relevance is plausible in the context (i.e., Mary’s need to see a doctor).

(17) *Mary has lived in London for five years. She has become ill.*

× Context A: Mary moved to London five years ago, and hasn’t left. During this time, she became ill only once, three years ago.

√ Context B: Londoners who have developed illnesses during the last five years are advised to go see their doctors, as their illnesses are likely due to some dangerous pollutants which were inadvertently released into the air.           (Portner 2003:463)

Current relevance is, however, not a component of the Bunun -in-. Example (18) shows that a sentence marked with -in- is not a felicitous response to a question regarding someone’s current state of recollection; this contrasts minimally with a felicitous sentence marked by the enclitic =in (to be discussed in Section 3). In other words, the infix -in- does not bear current relevance.

(18) Context: Do you remember that we studied Bunun together?

a. # Wa, s<in>ipungul saikin.
   INTJ <E.PST>AF.forget 1SG.NOM
   Intended for ‘I’ve forgotten.’

b. Wa, sipungul=in saikin.
   INTJ AF.forget=COS 1SG.NOM
   ‘I’ve forgotten.’

Another pragmatic effect that the English present perfect incurs is lifetime effects. Lifetime effects describe that the grammatical subject of a present perfect sentence must be alive at the utterance time for the present perfect sentence to be felicitous (see related properties such as repeatability or future possibility, whereby recurrence of the event in question need to be possible at the utterance time; e.g., McCawley 1971; Inoue 1979; Katz 2003; Portner 2003). An example is given in (19).

(19) # My late grandma has given birth to three children.

However, (20) clearly illustrates that no such pragmatic effect is observed with the Bunun -in-.

(20) Context: Talking about our late grandma.

Ta<in>us-uvaaz inaak nas-cinahudas tu ta~tau mas ’uvaaz.
AF.bear<E.PST>=child 1SG.POSS late-grandmother PRT HUM~three OBL child
‘My late grandma gave birth to three children.’

The English present perfect is also known to be associated with a recent past reading, but the Bunun -in- does not allow it. For instance, as shown by the contrast in (21), the discovery of the recent death of a pet cannot be expressed by the verb of dying mataz infixed by -in-; rather this is rendered by the same verb encliticized by =in.
(21)  Context: A kid interrupts: ‘‘Miss Savi, our pet rat has just died!’’ (Elicited based on ‘Miss Smith’s Bad Day’)
   a. # İmita  s<in>aipuk  tu  aluaz=a  m<in>ataz.
      İNCL.POSS RF<E.PST>raise LNK rat=DIST.NOM <E.PST>AF.die
     Intended for ‘‘The rat we keep has died.’’
     Consultant’s comment: “It’s not possible that it revives. People die once, except for Jesus.”
   b. İmita  s<in>aipuk  tu  aluaz=a  mataz=in.
      İNCL.POSS RF<E.PST>-raise LNK rat=DIST.NOM AF.die=COS
     ‘‘The rat we keep has died.’’

When marked by the present perfect, telic predicates allow for a result state reading, as exemplified in (22) (with the inferred result state in brackets). By contrast, the Bunun -in- is not felicitous for a result state reading; this is exemplified in (23), where the intended result state of the speaker’s wallet being at home is only expressible by =in.

(22)  a. I have opened the door. [The door is open.]
   b. I have arrived in Paris. [I am in Paris.]

(23)  Context: On your way to a store, you realized that you didn’t bring your wallet.
   a. # Ma-<I>kaunga  saikin  madas  inaak  tu  patsuian.
      AF<=>E.PST>leave.behind 1SG.NOM AF.carry 1SG.POSS LNK wallet
     Intended for ‘‘I’ve left my wallet at home.’’
     Consultant’s comment: “You are talking about your experience: for example, there was a time you
      left your wallet at home after you left the house”.
   b. Kaunga-an=in=ku  inaak  patsuian  sia  lumah.
      leave.behind-LF=COS=1SG.OBL 1SG.POSS wallet LOC house
     ‘‘I’ve left my wallet at home.’’

Universal perfect readings obtain in situations where an event has been going on since a past time and persists up to the utterance time. In English, universal perfect readings are only possible with stative predicates or predicates marked with the progressive, as illustrated by (24).

(24)  a. Mary has been angry since yesterday.
   b. John has been watching TV since 6 o’clock this morning.

In Bunun, however, the infix -in- does not allow these readings, with or without Ca-reduplication, as shown in (25)a-b; rather, the intended universal perfect reading uses Ca-reduplication without -in-.

(25)  Context: Another kid complains, “Miss Savi, he has been pulling my hair!” (Elicited based on ‘Miss Smith’s Bad Day’)
   a. # Maisi-kitngaab  hai,  ma-<I>damu  inaak  tu  hulbu.
      AF.be.from-begin TOP,  AF<=>E.PST>pull 1SG.POSS LNK hair
     ‘Since the beginning of class, he has been pulling my hair.’
   b. # Maisi-kitngaab  hai,  ma-<I>da∼damu  inaak  tu  hulbu.
      AF.be.from-begin TOP,  AF<=>E.PST>IPFV∼pull 1SG.POSS LNK hair
     ‘Since the beginning of class, he has been pulling my hair.’
Lastly, the fact that the infix -in- does not allow free reference times not only suggests that it is unlike an aspect (see also Section 2.1.1) but also that it is interpreted as relating reference time to utterance time. We expect that -in- does not behave as a past perfect in expressing an event anterior to a past reference time. This is borne out: the response to the question in (26) is intended to convey that the light-switching event occurs earlier than the time when the speaker reached home; in this case, =in rather than -in- is the correct rendition.

(26) Q: Masa ku-lumah kasu hai, maza sadu-an=su?
   when.PST go-house 2SG.NOM TOP what see-LF=2SG.OBL
   ‘When you reached home, what did you see?’

   A1: # Mai<si~>Singhal a dingki.
      <E.PST><IPFV~>AF.be.luminous NOM light
      Intended for ‘The light had been switched on.’

   A2: Mal<si~>singhal=in a dingki.
      <IPFV~>AF.be.luminous=COS NOM light
      ‘The light had been switched on.’

2.1.2.2 Experiential readings are existential
A remaining possibility is that -in- is a perfect aspect used exclusively for experiential readings. However, -in- in (7), (9)-(10), and (15)-(16) above does not have an experiential reading but a past tense reading instead. Specifically, experiential readings do not arise with a past-time adverb or in a past context. We argue that the reading of -in- is only existential; further evidence for the existential quantification of -in- is given in Section 2.3.

2.2 Interim summary
We have shown that the infix -in- exhibits properties that do not support a perfective analysis; especially, -in- co-occurs with Ca-reduplication (which only produces imperfective-like readings), yielding a past progressive or habitual reading. Also, the use of -in- is restricted to the past, and with stative predicates it gives a cessation inference. The simple past readings and cessation effects are unexpected for a pure perfective aspect but follow naturally from a past tense analysis. A rejection for -in- being a past tense marker is that -in- has salient experiential readings as found with the English perfect. However, a thorough examination of possible perfect properties shows that the infix -in- has nothing else in common with the English present perfect, and importantly, the experiential readings do not necessarily arise.

2.3 Evidence that -in- is an existential past tense
In this section, we aim to unify the experiential and simple past readings of the infix -in-. We argue that it is a past tense marker that encodes existential quantification over past times, following recent proposals in Chen et al. (2017) for Javanese and Atayal (see also Sharvit 2014; Thomas 2014; Mucha 2017). We offer two pieces of evidence. For one thing, -in- is infelicitous in deictic and anaphoric contexts, and for the other it shows scopal interactions with negation. These two properties are unexpected for a pronominal tense but fully compatible with an existential quantifier analysis; they, together with the experiential reading, strongly support that the infix -in- is an existential past tense marker.

2.3.1 Existential past vs. pronominal past
The existential past and pronominal past analysis are two competing approaches to the semantics of tense, mostly based on English data. An existential past tense is an operator encoding an existential quantifier over past times, often attributed to Prior (1967) (see Ogihara 1996; Kusumoto 2005; von Stechow 2009; a.o.). A
pronominal past tense is paralleled to a pronoun, which refers to a contextually salient past time (i.e., deictic uses) or a past time that has been established in the context or a narrative (i.e., anaphoric uses) (Partee 1973, 1984; Heim 1994; von Stechow 1995; Kratzer 1998; a.o.). While it has been debated whether the English past tense is existential or pronominal, past tenses in some languages have been argued to be unambiguously existential; for example, in Japanese (Sharvit 2014), Medumba (Bantu; Mucha 2017), Mbyá (Guaraní; Thomas 2014), Javanese and Atayal (Austronesian; Chen et al. 2017). We argue that the Bunun -in- closely resembles its cognate -in- in Atayal in exhibiting properties that are only explainable under an existential analysis.

2.3.2 Infelicity of -in- in deictic and anaphoric contexts
Unlike a pronominal past tense, an existential past tense is infelicitous in deictic or anaphoric contexts. The Bunun -in- behaves as predicted by an existential past. This is illustrated by (27), where the infix -in- is banned in reference to a salient contextual past time, namely the time before the speaker left the house.

(27)  Context: Driving on the highway after leaving the house, you realize (adapted from Partee 1973):
   a. I didn’t turn off the stove! (Partee 1973:602)
   b. Aa! Nii tu {#<in>ukud-an / sukud-an}=ku gasu.
      INTJ NEG LNK <E.PST>turn.off-LF / turn.off-LF=1SG.OBL gas
      ‘Oh! I didn’t turn off the gas.’

   Likewise, the Bunun -in- does not move the reference time forward in narratives; this is illustrated by (28), where a sequence of the watering event following the weeding event must be expressed by an unmarked verb. Using the same verb marked by -in- instead is infelicitous in this context and yields a back-shifting reading (i.e., the watering event occurs before the weeding event). Note that in contrast to the Bunun -in-, the English past tense is felicitously used in both deictic and anaphoric contexts.

(28)  Context: Describe what you saw Abus do this morning.
   a. Malabut saia ismuut at ma<suul saia hana.
      AF.remove 3SG.NOM grass CONJ <E.PST>AF.water 3SG.NOM flower
      Intended for ‘She weeded, and watered flowers.’
      (OK: ‘Having watered flowers, she weeded.’)
   b. Malabut saia ismuut at masuul saia hana.
      AF.remove 3SG.NOM grass CONJ AF.water 3SG.NOM flower
      ‘She weeded, and watered flowers.’

2.3.3 Interpretation of -in- under negation
Due to existential quantification, an existential past tense is expected to exhibit scopal interactions with operators like negation, whereas a pronominal past would not be able to do so. It turns out that the infix -in- shows scopal properties. Scoping -in- under negation yields a negative experiential reading (i.e. narrow-scope existential). For example, the negation nii preceding -in- in (29) gives rise to non-existence of the event in question (‘never’); a pronominal past tense would not yield such a reading (but one such that ‘I didn’t have breakfast at a contextually salient time’). It is unclear to us at this stage whether -in- can scope over the negation nii to yield an inverse scope reading.

---

22 Example (28) clearly shows that at least in narrative contexts, an -in- sentence does not freely alternate with a bare sentence. This thus falsifies Jeng’s (1999:462) claim that -in- sentences form a free variant with bare sentences.

23 A verbal predicate is negated by either the negator nii alone (as in (30) and (31)) or the negator plus the linker tu (as in (29)). The two patterns are conditioned by whether the negator is adjacent to the verb without intervening constituents like pronominal clitics or free forms (Huang and Shi 2016:159; Li 2018:471) and is irrelevant to the availability of the ‘never’ reading.
Context: Describing your life experience.

\[
\text{Nii tu } m<\text{in}>aum \ mas \ pinit'um'um. \\
\text{NEG LNK } <\text{E.PST}>\text{AF.eat} \ \text{OBL} \ \text{breakfast} \\
\text{‘I have never had breakfast.’}
\]

Notice that the co-occurrence of -in- and negation is not restricted to negative experiential readings. The temporal domain of existential quantification that -in- encodes does not have to encompass one’s entire lifetime: it can be restricted to a specific time span, as indicated by a temporal adverbial in (30) or contextually implied in the question and answer of (31). Both examples show the past time that -in- quantifies over is domain-restricted.

(30) Context: He is very sleepy now because…

\[
\text{Nii sanglabian } ma<i>\text{sabah.} \\
\text{NEG last.night } <\text{E.PST}>\text{AF.sleep} \\
\text{‘He didn’t sleep last night.’}
\]

(31) Context: Your friend is looking for his wallet.

Q: \[
<\text{E.PST}>\text{AF.see} \ \text{2SG.NOM} \ \text{1SG.POSS} \ \text{LNK} \ \text{wallet} \\
\text{‘Did you see my wallet?’}
\]

A: \[
\text{Nii saikin } s<\text{in}>\text{adu} \ \text{isuu} \ tu \ patsuian. \\
\text{NEG } \text{1SG.NOM} <\text{E.PST}>\text{AF.see} \ \text{2SG.POSS} \ \text{LNK} \ \text{wallet} \\
\text{‘I didn’t see your wallet.’}
\]

2.4 Analysis of -in-

We assign a domain-restricted quantificational semantics to -in-, following the precedents (Ogihara 1996; von Stechow 2009; Mucha 2017; Chen et al. 2017; a.o.). Essentially, -in- as a tense first adjoins to a domain restriction variable C, the result of which denotes a function that takes a predicate of times P and an evaluation time t, and asserts that a time t′ precedes t, at which P holds. The t argument will be saturated by the utterance time, tc.

\[
\left[ -\text{in-} \right]^{\mathcal{C}} = \lambda C, t. \lambda P. \lambda t. \exists t' \ [t' < t \land C(t') \land P(t')]
\]

The domain restriction variable ensures that the past times quantified over by -in- are restricted to a salient contextual time. This captures apparently varied readings of -in-: when there is an (implicitly or explicitly) established past reference time, -in- quantifies over the past time, and gives a past-tense reading; in the absence of a salient past reference time, the domain of the existential quantifier is naturally the span of one’s life, hence an experiential (perfect) reading.

What should be emphasized is that while the domain restriction of -in- resembles a pronoun referring to the reference time, -in- still encodes a quantifier, and as such it is not equivalent to a temporal anaphor. The function of -in- is not to directly pick up that interval but to quantify over a(n) (referred) interval; the difference is much like the one between a definite NP or pronoun and an indefinite member of a definite set in the nominal domain. This explains the unacceptability of -in- in the Partee’s stove context (see (27) above): although the context does provide a salient interval for the domain variable of -in-, what it calls for is a definite past tense rather than an indefinite one. By contrast, -in- is felicitous in the question of (31), which is similarly uttered with a contextually established salient time (i.e., a while before your friend’s search), because it concerns about the existence of the seeing event within that time.

Our analysis also reveals that -in- differs from morphologically tenseless sentences that are interpreted in the past (e.g., (28)b above) (cf. fn. 22). In other words, -in- is not a past tense optionally used to restrict the temporal interpretation of morphologically tenseless sentences, unlike what is claimed for in other languages (e.g., Bochnak 2016). Looking into the semantics of past tense markers hence provides an alternative to the seeming optionality in the overall tense system in Bunun. For future research, it is worth exploring tense of
3 The semantics of =in

This section turns to the enclitic =in. Departing from most of the previous literature, we argue that =in is not an aspectual marker but a discourse marker that induces a certain change of state. While the change of state of =in does not go unnoticed (e.g., Huang 1997:383; Su 2008:23; Chen 2009:20; Huang and Shi 2016:117), we show that it is pragmatically conditioned; for this reason, we argue that any aspectual analysis of =in is inadequate.

3.1 Evidence that =in is not a perfect aspect

As reviewed in Section 1.2, the enclitic =in has been treated as a perfect aspect based on either a universal perfect or anteriority reading. However, these claims are mostly based on data without any contexts. Sections 3.2 and 3.3 will show that =in always gives rise to a change of state, at a propositional or discourse level, and as a result, universal perfect readings (which involve a homogeneous property rather than a change of state) are only attested when the change of state is interpreted at a discourse level. In this subsection, we focus on evidence against an anteriority-based perfect analysis of =in as in Jeng (1999).

Sentences with =in have been argued to express perfect meanings based on out-of-the-blue contexts, where the reference time is the utterance time. Our examination of cases where =in occurs in a past or future reference time shows that the anteriority reading does not necessarily hold. With a past reference time introduced by a masa-clause, the matrix going event may precede or occur at same time with (but not follow) the reference time, as in (33). With a future reference time introduced by a mais-clause and marked by na= in the matrix clause, the marker =in commonly receives an imminent future reading, as in (34) (see also Huang 1997:383; Jeng 1999:476).

(33) **Masa** tauna-lumah saikin hai, *mu-daahn=in* saia.
    When.PST AF.reach-house 1SG.NOM TOP AF.go-road=COS 3SG.NOM
    ‘When I reached home, (s)he was leaving.’ / ‘When I reached home, (s)he had left.’

(34) **Mais** tauna-lumah saikin hai, *na=mu-daahn=in* saia.
    when.NPST AF.reach-house 1SG.NOM TOP FUT=AF.go-road=COS 3SG.NOM
    ‘When I reach home, (s)he will be leaving.’ / ≠ ‘When I reach home, (s)he will have left.’

The progressive reading in (33) and the imminent future reading in (34) clearly deviate a perfect aspect, which would give a past perfect reading with a past reference time or a future perfect reading with a future reference time. It is noteworthy that the imminent future reading is similarly attested with other types of predicate marked with =in, for example, an activity predicate in (35):

(35) Context: You come to your friend’s home when she’s busy. She says, “Please have a seat and wait...”
    **Na=ma-pa-tal isuh=in** saikin Aping=cia.
    FUT=AF-CAUS-shower=COS  SG.NOM Aping=DIST.OBL
    ‘I am about to give Aping a bath.’ / ≠ ‘I will have {started/ been in the process of} giving Aping a bath.’

3.2 Evidence that =in is not a perfective aspect

It has been debated whether =in yields a completable or continuous reading (Zeitoun et al. 1996; Jeng 1999; De Busser 2009), and the completable reading has led to a perfective analysis (De Busser 2009). However, Huang (1997) and Huang and Shi (2016) correctly point out that in Isbukun, the reading of =in (in out-of-the-blue context) varies with the type of predicate: completion readings with achievements, inceptive readings with activities, and inchoative readings with statives; some of their examples (translated from Mandarin) are given here.
(36) $\text{Mataz} = \text{in}$ a k\text{\textlangle in\textrangle} alat mas Subali =cia a 'asu=a. \text{[Achievement]}
\begin{footnotesize}
AF.die=\text{COS} \quad \text{NOM } <\text{\textlangle E.PST\textrangle}\text{AF.bite } \text{OBL} \quad \text{Subali=\text{DIST.OBL NOM} } \text{dog}=\text{DIST.NOM}
\end{footnotesize}
‘That dog that bit Subali died.’ (Huang and Shi 2016:223)

(37) $\text{Tangis} = \text{in}$ a 'isaicia a 'uvaaz. \text{[Activity]}
\begin{footnotesize}
AF.cry=\text{COS} \quad \text{NOM 3SG.POSS LNK} \quad \text{child}
\end{footnotesize}
‘His child cried (and is still crying).’ (Huang and Shi 2016:118)

(38) $\text{Ma-diav} = \text{in}$ a bunbun=a. \text{[State]}
\begin{footnotesize}
AF-yellow=\text{COS} \quad \text{NOM} \quad \text{banana}=\text{DIST.NOM}
\end{footnotesize}
‘Those bananas became yellow.’ (Huang and Shi 2016:118)

It should be noted that statives marked with =\text{in} are only interpreted as inchoative; for instance, placing (39) in a context that targets a homogeneous state is rejected:

(39) Context: You are teaching kids the color. You say, “Look. Those bananas are yellow.”
\begin{footnotesize}
#\text{Ma-diav} = \text{in} \quad a \quad \text{bunbun=a}. \text{[State]}
\end{footnotesize}
Intended for ‘Those bananas are yellow.’

To complete the picture, we include accomplishment predicates, which are not covered in the literature. Much like activities, accomplishments marked with =\text{in} have inceptive readings; (40) shows that it is felicitous to continue the =\text{in} sentence with the statement that the accomplishment event has not culminated.

(40) $\text{Ka-lumah} = \text{in}$ saikin tu dusa, ka-nii=ang ka-nahtung-an. \text{[State]}
\begin{footnotesize}
AF.build-house=\text{COS} \quad 1\text{SG.NOM PRT} \quad \text{two build-NEG=still build-finish-LF}
\end{footnotesize}
‘I built two houses, but I haven’t finished (them) yet.’

It is clear from these above readings that =\text{in} does not always induce a completive reading with telic events (i.e., accomplishments and achievements) as a perfective aspect analysis would predict; notice that we adopt a notion of completion which refers to the final points/culmination of telic events (cf. Smith 1997). Instead, the resulting reading varies depending on the durativity of events: completion for achievements and non-completion/continuation for activities and accomplishments.

(41) Context: He used to be fat, and he is still fat now.
\begin{footnotesize}
Mai-si-\text{kauma} 'ikit saia masmuh(*=\text{in}).
\end{footnotesize}
‘He has been fat since he was young.’

(42) Context: Another kid complains, “Miss Savi, he has been pulling my hair!” (Elicited based on ‘Miss Smith’s Bad Day’)
\begin{footnotesize}
Mai-si-\text{kinga}ab hai, {\text{\textlangle 1SG.POSS LNK hair}} {\text{\textlangle =\text{in} \quad \text{bunbun=a}.}} \text{inak tu hulbu.}
\end{footnotesize}
‘Since the beginning of class, he has been pulling my hair.’

24 While the fact that accomplishment predicates marked with =\text{in} are not culminated may be paralleled with non-culminating perfectives in other languages (see e.g., Singh 1998; Koenig and Muansuwan 2000; Soh and Kuo 2005; Bar-el 2005; Koenig and Chief 2008; a.o.), a change-of-state proposal better captures all the readings of =\text{in}.
In the next subsection, we present a failed attempt at analyzing \(=in\) as change-of-state/inchoative aspect; we discuss the reason why such an aspectual analysis is unattainable.

### 3.3 Evidence that \(=in\) is not an inchoative aspect

A remaining possible analysis of Bunun \(=in\) is that it is an aspect marking inchoativity/inception. Such an analysis has been proposed for the Samoan ‘uo (Hohaus 2016).\(^{25}\) According to Hohaus (2016:101), ‘ua is used with stative predicates to convey that “a change of state has taken place and that the respective state did not hold before the evaluation time”. As a result, it cannot co-occur with an individual-level state, as in (43), and is unacceptable in contexts where the described event has begun before the evaluation time, as in (44).

#### (43)
Context: Providing a character description of your friend John:

# ‘O loane ‘ua sau mai Egelagi.

FOC John INCH come from England

Intended for ‘John, he is from England.’ (Hohaus 2016:101)

#### (44)
Context: Your grandmother broke her arm three weeks ago and has been in a lot of pain ever since.

Today, her doctor called her to ask: “How are you feeling today?” She replied:

# ‘Ua tīgā lo‘u lima.

INCH painful my arm

‘My arm is (now) painful.’ (Hohaus 2016:102)

Hohaus argues that ‘ua requires the evaluation time (similar to the reference time, whose value is given by tense) be the initial sub-interval of the running time of the eventuality. In other words, ‘ua specifies that an initial sub-interval of the event coincides with the reference time. In what follows, we discuss advantages and disadvantages of applying this analysis to \(=in\) and eventually conclude that it is undesirable.

Firstly, the Bunun \(=in\) exhibits the initiality requirement observed for the Samoan ‘uo. In cases where the reference time refers to the utterance time, \(=in\) requires that the beginning runtime of the event be very adjacent to the utterance time. This is evidenced by its incompatibility with past-time adverbials: as shown in (45), while \(=in\) is compatible with laupakadau ‘now’, it cannot co-occur with laupang ‘just, very recently’ or any past-time adverbs in out-of-the-blue contexts.\(^{26}\)

#### (45)
a. Masabah=\(=in\) saia laupakadau.

AF.sleep=COS 3SG.NOM now

‘(S)he is falling asleep now.’

b. * Laupang saia masabah=\(=in\).\(^{27}\)

just 3SG.NOM AF.sleep=COS

Intended for ‘(S)he has just fallen asleep.’

c. * Masabah=\(=in\) saia \{takna / sangan\}.

AF.sleep=COS 3SG.NOM yesterday / a.while.ago

Intended for ‘(S)he \{slept/ was sleeping\} \{yesterday/just now\}.’

An inchoative aspect proposal could explain why the reading of \(=in\) varies with a different reference time (Section 3.1). When the reference time denotes an instantaneous time interval—for example the utterance time or a punctual reference time—an initial sub-interval of an event, despite spanning a short time, inevitably

---

\(^{25}\) It is also comparable to the Niuean \(kua\) (Matthewson et al. 2015) without utilizing perfect semantics.

\(^{26}\) Jeng (1999:477) notes that “the Bunun perfect can co-occur with the present, past, and future tenses”; however, all his examples where \(=in\) and past-time adverbs co-occur also involve the use of -\(in\)-. As we show in (45)b-c, \(=in\) alone is not compatible with past-time adverbs.

\(^{27}\) Unlike other temporal expressions, laupang behaves more like a verb than an adverbial because it can host nominative person clitics (Li 2018:225) or immediately precede nominative NPs.
includes the reference time, hence an anterior reading. That is, an anterior reading parallel to a past tense or a past perfect is simply a pragmatic effect of matching the beginning of an event and the utterance time. With the adjacency requirement, we also expect a simultaneous/on-going reading. Similarly, when the reference time is in the future, the described event must begin at the future time, hence an imminent future reading. Moreover, the aspectual analysis also captures the different interpretive effects varying with the durativity of events (Section 3.2). Since \( =in \) only concerns the initial status of events, process events (activities and accomplishments) naturally continue at the reference time whereas achievements are themselves instantaneous so that the entire event must be no later than the instantaneous reference time, thus yielding a completion reading (i.e., they begin and complete at the same time at the reference time).

However, the inchoative aspect analysis does not correctly account for the following fact: \( =in \) can induce a separate non-inchoative reading given an appropriate context. Recall that, as shown in (41)-(42) above, the inchoative reading of \( =in \) is incompatible with a time interval stretched from the past up to the utterance time. However, the sentence with \( =in \) in (46), which forms a near-minimal pair with (41), has an unambiguous universal perfect reading. What it differs from (41) is that \( =in \) in (46) indicates that the proposition—a third person referred to has been fat—is counter to the interlocutors’ expectation, specifically that of the addressee/listener.

(46) Context: You doubt he was not fat when he was little, but I think:

\[
\text{Maisi-kauma’ikit saia hai, nau tu massmuh=\textit{in}.}
\]

\begin{verbatim}
AF.be.from-little 3SG.NOM TOP should LNK AF.fat=COS
\end{verbatim}

‘Since he was little, he has been already fat for sure.’

The fact that the “change-of-state” and the “contrary-to-expectation” interpretations are expressed by the same form is reminiscent of the Mandarin sentence-final particle \( le \), which has been analyzed as a marker that induces a change within or across propositions assumed by interlocutors (Soh 2009). The similarity suggests that \( =in \) is more of a discourse-level change-of-state marker, which explains why it does not consistently contribute an inchoative reading as an aspectual marker would do. Future investigations will be benefited from comparing \( =in \) and “change”-inducing elements cross-linguistically, such as English \textit{already} (Löbner 1989; Mittwoch 1993; Vander Klok and Matthewson 2015; a.o.) and so-called iamitive markers in many Southeast Asian languages (Olsson 2013; Dahl and Wälchli 2016), and from situating \( =in \) in the studies of Common Ground (Stalnaker 1999, 2002).

Building the precise semantics of \( =in \) will also require a clear understanding of how \( =in \) is used in combination with other aspectual/temporal markers, including the existential past tense \( -in- \), and of whether a contrary-to-expectation reading is equally available in the combinations. We leave a formal analysis of \( =in \) and establishing the overall architecture of tense and aspect in Bunun for another occasion.

4 Concluding remarks

In this last section, we summarize our main findings and discuss theoretical and typological implications of our proposal for studying similar phenomena in other languages, especially other Formosan languages.

4.1 Summary of the findings

By presenting new data, we have provided an alternative account for the semantics of \(-in-\) and \( =in \) in Bunun. We argue that the two markers differ in how they express anteriority. The infix \( -in- \) is a past tense, and lexically denotes existence of events in the past. The domain-restricted existential quantification of \( -in- \) explains why its reading is sometimes comparable to an experiential-perfect reading and sometimes a past-tense reading. By contrast, the enclitic \( =in \) induces a change of state, either at the level of propositional meaning or discourse. At the propositional level, the anteriority effect of \( =in \) arises possibly due to coincidence of the inception of events and the reference time. We also suggest the direction of future work on the non-inchoative reading of \( =in \). The consequence of our findings for the two homophonous markers refutes the claim that Bunun has both perfect and perfective category (pace Zeitoun et al. 1996).
4.2 Theoretical and typological implications

The proposed semantics for the two anteriority morphemes in Bunun contributes to the theory of temporal semantics and the idea of decomposing temporal categories. The infix -in- supports the existence of existential past tense (Sharvit 2014; Thomas 2014; Mucha 2017; Chen et al. 2017) and clearly shows that the semantics of tense can be empirically distinguished but not merely reflects notational variants. Even within languages that are argued to have existential past tenses, the existential past tenses may not behave exactly the same; for example, the Bunun =in differs from the existential past markers in Medumba, whose quantificational domain is lexically restricted (Mucha 2017). The enclitic =in shows not only that properties partly similar to perfective or perfect can be unified by change of state but also that the change of state need not be encoded in an aspect.

Typologically, our finding suggests the need for a re-assessment of temporal/aspectual markers that bear similarities to these two markers in other Austronesian languages, especially other Formosan languages. First, our analysis of the Bunun infix -in- has implications for widespread reflexes of PAn *-in- in Western Austronesian. These reflexes are generally described as marking “past tense or completive aspect” (Blust 1998:347) or “perfective aspect” (Blust 2013:385). However, a perfective analysis for -in-, at least in Bunun (this work) and Atayal (Chen et al. 2017), is unattainable. Consequently, it is questionable whether a perfective analysis equally applies to other reflexes.

Second, a perfect or perfective analysis for =in in Bunun has difficulty in unifying all of its possible readings. Our proposal that =in is a dedicated change-of-state marker could be possibly extended to similar markers in other Formosan languages, many of which have also been described as perfective markers. Among Formosan languages, change of state is expressed either by sentence-final particles, as in Atayal (Rau 1992:158), Seediq (Sung 2016:89), and Saisiyat (Huang 2003:98) or by head-adjacent clitics, as in Bunun, and all the other Formosan languages, including Amis (Wu 2006:123), Kavalan (Hsieh 2016:79), Thao (Wang 2004:216), Puyuma (Teng 2008:32), Paiwan (Li 2010:37), Rukai (Chen 2008:179), and Tsou (Pan 2007:42). Given the fact that they are similarly reported to describe the beginning of an eventuality, it is expected that they share more functional affinities with the Bunun =in than perfective/perfect aspect markers.28

Our study also bears a methodological point. The semantics of -in- and =in is uncovered through examination against the properties of temporal categories. This shows that in order to approach semantic nuances of temporal/aspectual markers in languages, a targeted and hypothesis-driven study is desired.

References


---

28 For instance, a change-of-state analysis may be applied to account for why Kavalan =ti incurs two different readings on the same motion verb: an anterior reading on the bare form, e.g. maseq=ti (AF.arrive=COS) ‘to have arrived’, but an inchoative reading on the future form, e.g. qaseq=p=ti (arrive=FUT=COS) ‘be about to arrive’ (Hsieh 2016:80).


Li, Chao-Lin. 2010. The syntax and semantics of eventuality in Paiwan and Saaroa. PhD diss., National Tsing Hua University.


Lin, Kuo-Chiao. 2010. The infix -i(n)- in Takituduh Bunun and the syntax of inner aspect. MA thesis, National Taiwan University.


MORE ON KANAKANAVU WORD-LEVEL PROSODY:

CYCLIC AND POSTCYCLIC PROCESSES

Yi-Yang Cheng
University of California, Santa Barbara
ycheng@ucsb.edu

Abstract
The present paper introduces additional data and observations regarding patterns of word-level prosodic prominence in Kanakanavu. An integrated analysis couched in Lexical Phonology is proposed in which two types of prominence are distinguished. H-prominence, which is characterized by a high pitch target, is attributed to a H tone that is assigned cyclically and interacts with other cyclic and postcyclic rules in the derivation. HL-prominence, which is characterized by a falling pitch contour, is attributed to a HL tone that is underlyingly specified in certain morphologically simplex forms and may or may not block the cyclic assignment of the H tone. The proposed analysis takes into account a broader consideration of rich morphophonology in Kanakanavu, which sheds light on both the nature and types of word-prosodic prominence in the language and calls for a rethinking of the label “stress” that has been widely employed in the literature.

Keywords: Kanakanavu, word-level prosody, morpho-phonology, prosodic typology

ISO 639-3 codes: xnb

1 Kanakanavu word-level prosody in the literature
Kanakanavu is a critically endangered Formosan language spoken in the Namasia District of Kaohsiung, southern Taiwan. Although the total number of people who identify as Kanakanavu is currently around 350, there are now fewer than 10 speakers who are fluent in the language. Most Kanakanavu people are found to be fluent in Mandarin and/or Bunun, two languages they have been in close contact with (Liu et al. 2015).

Despite the endangered status of Kanakanavu, there has been a body of literature in Formosan linguistics built on describing several aspects of its linguistic structure, and one of the main threads of discussion has been concerned with the nature of word-level prosodic prominence.

---

1 The first-hand data presented in this paper were collected in collaboration with the Kanakanavu community at the Namasia District of Kaohsiung, Taiwan, with special assistance from community language activists Mr. 'Angai Ka'angaiana (翁博學) and Mr. Pani Kanapaniana (孔岳中) during the summers of 2016 and 2017. Trips to the community and the data collection sessions were parts of an ongoing community-based documentation project to record Kanakanavu language use, which had been partially funded by the UCSB Chancellor’s Fellowship and the UCSB Linguistics Department Summer Research Grant. I greatly thank the Kanakanavu people for supporting my participation in this long-term collaborative work.

2 This paper is a crystallization of two term papers written under the supervision of Dr. Matthew Gordon and Dr. Eric Campbell, respectively. I thank them for all the invaluable intellectual resources they offered me in developing ideas at the earliest stages of exploring the topic. Some modifications have been made since the paper was presented at AFLA 25, Academia Sinica, Taipei, and they were inspired by comments offered by Dr. Hui-chuan J. Huang, Dr. Elizabeth Zeitoun, and other audience members at the conference. Additional modifications were made during the revision process thanks to the extremely helpful comments provided by two anonymous reviewers. Without them this paper would have been lacking in detail, depth and comprehensiveness. Finally, special thanks go to Sally Chen and Hsuan-ju Chen, who have been the pioneers in re-examining Kanakanavu prominence, and have been engaging in fruitful discussions with me as this paper develops. I remain solely responsible for all the errors and mistakes in the paper.
In the earliest studies on Kanakanavu phonology, word-level prosodic prominence was mainly described based on impressionistic observations.\(^3\) It has been commonly referred to as stress, but the relevance of pitch in its identification has also been repetitively recognized. Yan (1964:138-139), for example, explicitly mentions the close relationship between stress and pitch/tone in the language but points out that more detailed research is still required. In both M. Sung (1966:799) and Tsuchida (2003:10), the “stressed” syllables are also described as “coinciding with” or “accompanying” a high pitch.

Another observation made by early phonological studies on Kanakanavu is that word-level prosodic prominence shows variation in the position in a word on which it is realized. In Yan (1964:138-139), it is mentioned that the most commonly found positions are the penultimate and antepenultimate syllables of a word, whereas M. Sung (1966) maintains that those are the only positions where prosodic prominence can be found. Additionally, Tsuchida (2003) posits prosodic prominence in the language to be “phonemically distinctive” with the caveat that “the functional load is very low” (ibid. 10-11).\(^4\)

These two main themes that have emerged in early studies on Kanakanavu word-level prosodic prominence—its acoustic properties and its position/phonological properties—have been picked up by two recent studies that contribute to more detailed analyses from empirical, typological, and theoretical perspectives.

As the first study to employ the neutral term “word-level prominence” instead of “stress” to refer to prosodic prominence in Kanakanavu, S. Chen (2016) places Kanakanavu in a category different from other Formosan languages such as Paiwan and Saisiyat with regard to both the position and the phonetic cue of word prosodic prominence. Whereas Paiwan (C. Chen 2009) and Saisiyat (Chiang & Chiang 2005) have been described as “pitch-accent” languages, prominence in the two languages is found to occur in fixed positions, with pitch being the only phonetic cue of its realization. The acoustic study conducted by S. Chen (2016), however, shows that Kanakanavu not only exhibits more variability in where prominence may be realized—including the final, antepenultimate, and (the most commonly seen) penultimate syllables—but both pitch and vowel duration are found to serve as the most robust phonetic cues for prominence. In this regard, S. Chen’s (2016) results do in a way confirm the earlier impressionistic observations on the salient role played by pitch in Kanakanavu word prosodic prominence. Moreover, it also contributes to offering a preliminary typological picture in setting Kanakanavu apart from other Formosan languages such as Paiwan and Saisiyat, as it emphasizes the role vowel duration plays along with pitch in Kanakanavu, which had escaped scrutiny in earlier studies.

The other recent study that continues discussion on Kanakanavu prosodic prominence is H. Chen (2016), which is the first study to propose an account for the varying positions of prominence in the language. In this study, the realization of prosodic prominence is deemed “mostly predictable” (ibid. 83) as it can be attributed to (i) regular prominence realization on the underlying penultimate mora of a word, and to (ii) phonological factors that result in deviations from the regular penultimate position.

To begin with, what is referred to as “regular stress”\(^5\) is found on the penultimate syllable in words that have simple (CV)(CV)CVCV syllable structures, where the penultimate mora consistently coincides with the penultimate syllable. This can be seen in Table 1.

---

3 This was because description of the segmental aspects of Kanakanavu phonology had been one of the main focuses.

4 Tsuchida (2003:11) provides the following two minimal pairs that show a contrast between prominence on the penultimate syllable and that on the antepenultimate syllable, as can be seen below. H. Chen (2016:83), however, elicited the same data to verify the distribution of stress, but “[did] not see stress as phonemically distinctive”.

(i) nipapaciʔira ‘show (AV.PFV)’
(ii) nipapaciʔira ‘look at each other (AV.PFV)’
(iii) apaikiri ‘to let do tightly and compactly’
(iv) apaikiri ‘to let hold’

    cf. apaciʔira ‘show’
    cf. arupaciʔira ‘look at each other’
    cf. maikiri ‘to do tightly and compactly’
    cf. umikiri ‘hold’

(adapted from Tsuchida 2003:11)

5 Note that although H. Chen (2016) cites Sung (1966) in emphasizing the role pitch plays in Kanakanavu prominence, the term “stress” is still used. Regardless, the term “prominence” will be used when reviewing her study in this paper for coherence and consistency.
Table 1: “Regular” penultimate prominence in Kanakanavu words with (CV)(CV)CVCV syllable structure (adapted from H. Chen 2016:84-85)

<table>
<thead>
<tr>
<th>CVCV</th>
<th>CVCVCV</th>
<th>CVCV.CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>tī:ki ‘ear’</td>
<td>kakitsu ‘man’s basket’</td>
<td>talikúka ‘chicken’</td>
</tr>
<tr>
<td>?u.ru ‘cooked rice’</td>
<td>?enésa ‘there’</td>
<td>tepenáñe ‘bird’</td>
</tr>
<tr>
<td>vá:tu ‘stone’</td>
<td>saʔōʔo ‘delicious’</td>
<td>ʔatsipuña ‘a little’</td>
</tr>
<tr>
<td>pāːji ‘pestle’</td>
<td>tanáli ‘peanut’</td>
<td>ʔumítári ‘wait (AV)’</td>
</tr>
</tbody>
</table>

One phonological factor that productively results in deviation from the regular penultimate prominence pattern is syllable weight in word-final position. Words that contain underlyingly final heavy syllables, including native and loan words, are found to have prominence realized on the final syllable, which contains the penultimate mora in the underlying representation: 7

(1) CVCervoir (adapted from H. Chen 2016:85-86)
   a. punéː /punai/ ‘pigeon’
   b. kuléː /kulai/ ‘worm’
   c. riŋéː /riŋai/ ‘trap’
   d. vunéː /vunai/ ‘snake’
   e. tsanáː /tsanaa/ ‘field (wet rice)’
   f. tsanáː /tsaŋaa/ ‘green onion’ 8

Deviations from the regular penultimate position may also result from final heavy syllables created through affixation. For example, in words that end with open syllables, suffixation involving morphemes consisting of an onsetless syllable (e.g. the third person singular possessive -in) would result in heavy final syllables due to vowel coalescence, as can be seen in (2) 9 below. Here, prominence is found to be realized on the final syllable.

(2) CVCVCVCV
   a. punéː /punai/ ‘pigeon’
   b. kuléː /kulai/ ‘worm’
   c. riŋéː /riŋai/ ‘trap’
   d. vunéː /vunai/ ‘snake’
   e. tsanáː /tsanaa/ ‘field (wet rice)’
   f. tsanáː /tsaŋaa/ ‘green onion’ 8

Deviations from the regular penultimate position may also result from final heavy syllables created through affixation. For example, in words that end with open syllables, suffixation involving morphemes consisting of an onsetless syllable (e.g. the third person singular possessive -in) would result in heavy final syllables due to vowel coalescence, as can be seen in (2) 9 below. Here, prominence is found to be realized on the final syllable.

---

6 The lengthened vowels in disyllabic words in this table are proposed by H. Chen (2016) to be a result of a minimal trimoraic constraint in Kanakanavu words. Because the author of this paper remains neutral as to whether vowel length is realized phonetically in Kanakanavu, first hand data presented in this paper do not involve transcription of vowel length. However, there is evidence indicating the existence of underlyingly heavy syllables, especially those created through suffixation, their phonemic representation will be provided along with phonetic data.

7 H. Chen (2016:86-87) also analyzes final CVN syllables to be heavy, resulting in word-final prominence, as illustrated in the cited data below:
   (i) makúŋ ‘cold’
   (ii) morán ‘help (AV)’ (adapted from H. Chen 2016:87)

However, this analysis does not seem to hold due to the presence of counterexamples in the author’s fieldnotes, as seen below, and because words with final prominence-bearing CVN syllables can generally be accounted for by the productive process of Post-nasal high vowel deletion, which will be reviewed in (3) (this section) and discussed in more detail in (7-8) in Section 2.1.

   (iii) vóʔin ‘eye’
   (iv) tsakíɾan ‘river’

In addition, words that contain antepenultimate long vowels are also analyzed in H. Chen (2016:88-89) as resulting in antepenultimate prominence, as can be seen in the cited data below. Because the author of this paper does not assume vowel length in Kanakanavu (as mentioned in Footnote 6 above), these cases are tentatively treated as involving extrametrical echo vowels, like the words with antepenultimate prominence in (4).

   (i) kóːtsapa /kuátsapa/ ‘utensils’
   (ii) váːntuku /vanítuku/ ‘money’ (adapted from H. Chen 2016:89)

8 Glossing abbreviations used in this paper follow the Leipzig Glossing Rules. Additional abbreviations not included in the Leipzig Glossing Rules are as follows: AV, actor/agent voice; PV, patient voice; NIND, non-indicative; MIMP, mild imperative; STAT, stative.
Another phonological factor at play is the underlying statuses of word-final vowels. On the one hand, words involving underlying word-final high vowels (/i/, /i/, and /u/) that occur after nasal consonants are found to have prominence located on the final syllable due to consistent loss of these vowels in surface forms. As the underlying penultimate syllable—which carries the underlying penultimate mora—surfaces in word-final position, this results in prominence realized on the final syllable with a nasal coda:10

(3) CVCVCVN (adapted from H. Chen 2016:88)

| a. | makasiːn  /makasini/  ‘now’ |
| b. | miʔináːn  /miʔinani/  ‘before’ |
| c. | mitiviːŋ  /mitivini/  ‘hide oneself (AV)’ |

Words that contain final echo vowels,11 on the other hand, regularly involve prominence being realized on the antepenultimate syllable. This is attributed to the extrametrical status of these vowels in mora counting for prominence realization. Due to the extrametrical status of the final echo vowels, the underlying antepenultimate syllable carries the penultimate mora, hence its being the position in which prominence is realized:

(4) CVCVC<V> (adapted from H. Chen 2016:89)

| a. | vutúkuru  /vutukur<u>/  ‘fish’ |
| b. | ʔaŋírisi  /ʔaŋiris<i>/  ‘pork’ |
| c. | ʔitúmuru  /ʔitumur<u>/  ‘a lot’ |

The three attested positions of prominence realization—the antepenultimate, penultimate, and final syllables—as reported in S. Chen (2016), can then be complemented by H. Chen’s (2016) analysis, where different phonological factors come into play. This is summarized as in Table 2 below.

### Table 2: Prominence positions and interactions with phonological factors at play (adapted and summarized from H. Chen 201612)

<table>
<thead>
<tr>
<th>Factors</th>
<th>Antepenultimate</th>
<th>Penultimate</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular prominence on penultimate mora</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syllable weight (underlying final FR)</td>
<td>✓ (antepenult. heavy syll.)</td>
<td></td>
<td>✓ (final heavy syll.)</td>
</tr>
<tr>
<td>Syllable weight (final VV from affixation)</td>
<td></td>
<td>✓ (final heavy syll. from suffixation)</td>
<td></td>
</tr>
<tr>
<td>Final post-nasal high vowel loss</td>
<td></td>
<td>✓ (final CVN syll.)</td>
<td></td>
</tr>
<tr>
<td>Extrametricality of echo vowels</td>
<td>✓ (final echo vowels)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

10 These vowels are treated as echo vowels in H. Chen (2016:88), but they do not seem to be extrametrical, compared to those that are in (4). See Section 2.1 for more data showing how the underlyingly word-final vowels in (3) behave differently from other word-final vowels in Kanakanavu when morphology is considered.

11 Angle brackets here are the author’s own way of representing the extrametricality of final echo vowels.

12 H. Chen (2016) also examines other dimensions of prominence in Kanakanavu, including interactions with prefixes and clitics, prominence at the phrasal level, the possibility of “secondary stress”, etc. Due to space limitations, only data relevant to discussion in this paper have been reviewed.
Building on the contributions made by previous studies, the present paper aims to present additional data and observations on word-prosodic prominence in Kanakanavu. In light of the robust phonetic findings in S. Chen (2016), this study adopts the more neutral term “prominence” when referring to word prosodic prominence in the language throughout the paper. In addition, we agree with H. Chen (2016) in deeming that the location of prominence is predictable, and that there are divergences from the regular penultimate pattern that can be related to interaction with other phonological processes or the underlying representation of the forms in question.

To complement the results from previous studies, however, this study will incorporate a broader consideration of rich Kanakanavu morpho-phonology into an integrated analysis of word-prosodic prominence in the language. To achieve this goal, this paper will (i) further explore morphological factors at play in influencing prominence position, and (ii) argue for the existence of a second type of prosodic prominence that has not yet been discussed in the literature. The additional data will lead to proposal of an analysis of prosodic prominence in Kanakanavu couched in Lexical Phonology, in which the different prominence positions are analyzed as resulting from interactions between underlying representations and various cyclic and postcyclic rules in the derivation of words.

This paper is organized as follows. Section 2 will introduce the role that morphology plays in Kanakanavu word-level prominence by exploring how different underlying statuses of word-final vowels may interact with suffixation in giving rise to various prominence realization patterns. Section 3 will further demonstrate that certain morphemes in the language are specified as prominence-attracting or prominence-repelling, and that such specifications account for prominence realization patterns that would fail to be predicted if only phonological factors are taken into consideration. In Section 4, the prominence realization patterns discussed in the literature and in the preceding sections will be shown to involve only one (to be referred to as H-prominence) of the two types of prominence that can be identified in Kanakanavu. The other type of prominence—to be termed HL-prominence—will be introduced and discussed regarding its phonetic realization, invariable penultimate position and lack of interaction with morphology. An integrated analysis of Kanakanavu word-level prominence couched in Lexical Phonology will be proposed in Section 5, and Section 6 offers the summary and conclusion of this paper, where a property-driven approach to characterizing/typologizing Kanakanavu word-level prominence will be entertained under the current analysis.

2 Interactions of word-final vowels with suffixation in Kanakanavu

This section aims to introduce the role Kanakanavu morphology plays by investigating its interactions with underlying representation of words. As has been mentioned in Section 1 above, the underlying statuses of word-final vowels may interact with prominence realization in different ways. By further examining how suffixation interacts with word-final vowels in leading to different prominence positions, the discussion that follows will demonstrate that at least four types of word-final vowels are crucially distinguished in their underlying representation: (i) regular vowels, (ii) post-nasal high vowels, (iii) extrametrical echo vowels, and (iv) epenthetic echo vowels.

2.1 Final regular vowels and post-nasal high vowels

In Kanakanavu, all suffixes are found to begin with a vowel (cf. L. Sung 2018:28). Concatenation of word-final vowels and the initial vowel of a suffix will therefore result in final heavy syllables in the underlying representation of the suffixed form, with the originally word-final vowels carrying the penultimate mora. The suffixed forms then surface with prominence realized in word-final position, where different hiatus resolution patterns can be observed in leading to different surface syllable types. As can be seen in (5) below, the involvement of the actor voice imperative suffix -a may lead the stem-final vowel to undergo glide formation (5a), or coalescence (5b):

Due to the focus of this paper on prominence realization in Kanakanavu, which, as will be demonstrated, is largely dependent on moraic segments, obstruents in data from the author’s fieldnotes are transcribed phonemically across the board to avoid convolution in data discussion. However, it should be noted that dental/alveolar /s/ and /ʦ/ undergo palatalization before a high front vowel /i/: /s/ → [ɕ] / __ i, /ʦ/ → [ʨ] / __ i.
Vowel coalescence is also observed when the third person singular possessive suffix -in is involved, with the final surface vowel [e] resulting from the coalescence of stem-final /a/ and suffix-initial /i/:

(6) /CVC-VN/ → [CVCVN]
   a. ʦínén /sína-in/  ‘mother (3SG.POSS)’  cf. ʦína  ‘mother’
   b. ʨínɛn /kína-in/  ‘food (3SG.POSS)’  cf. ʨína  ‘food’

Words that involve underlying post-nasal high vowels in final position, in contrast, do not show final heavy syllables in the underlying representation when suffixation is involved. This is attributed to a productive process—to be referred to as Post-nasal high vowel deletion in this paper—whereby an underlying final high vowel undergoes deletion when occurring after a nasal consonant. Post-nasal high vowel deletion applies both in unsuffixed forms, where the underlying penultimate syllable on which prominence is realized surfaces as the final syllable, as in (7), and in suffixed forms, where the underlying penultimate syllable also surfaces as the penultimate syllable, as in (8).

(7) /CVCVCVN[+hi]/ → [CVCVCVN]
   a.  pusuʔán /pusuʔani/  ‘put (AV)’
   b.  mitivি /mitivи/  ‘hide oneself (AV)’

(8) /CVCVCVN-V(N)/ → [CVCVCVN(N)]
   a.  pusuʔана /pusuʔan-a/  ‘put (AV.IMP)’
   b.  mitivия /mitivи-a/  ‘hide oneself (AV.IMP)’
   c.  pusuʔанин /pusuʔanin/  ‘put (PV)’
   d.  mitivиян /mitivин/  ‘hide oneself (PV)’

In this regard, prominence realization in suffixed words in Kanakanavu can be predicted based on whether stem-final vowels undergo deletion. In suffixed forms that show underlingly final heavy syllables, final prominence results from it falling on the underlying penultimate mora and surfacing in word-final position through different hiatus resolution patterns. In suffixed forms that do not show underlying heavy syllables due to Post-nasal high vowel deletion, penultimate prominence results from the underlying penultimate mora corresponding to the surface penultimate syllable.

### 2.2 Extrametrical and epenthetic echo vowels

In words involving final echo vowels, a similar prominence realization pattern can be observed, where underlying final heavy syllables in suffixed forms may or may not obtain depending on the status of the echo vowels. As proposed in H. Chen (2016), extrametrical echo vowels are present in the underlying representation of words, leading to prominence being consistently realized on the antepenultimate syllable. Data cited from H. Chen (2016) has been presented in (4) above, and (9) below show further data to be discussed:

(9) CVCVC<υ>
   a.  pakituru /pakitur<υ>/  ‘promise’
   b.  tumatuturu /tumatutur<υ>/  ‘tell (AV)’
   c.  vantuku /vantuk<υ>/  ‘money’
   d.  vutukuru /vutukur<υ>/  ‘fish’

---

14 This is included in one of the processes referred to as “vowel weakening” in Tsuchida (2003:8).
Interestingly, the extrametrical status of these echo vowels appears to be irrelevant to prominence realization in suffixed forms, as they are found to form underlying final heavy syllables due to concatenation with the initial vowel of the suffix. As can be seen in (10) below, the same prominence realization pattern found in (5) and (6) (where regular stem-final vowels form underlying heavy syllables with suffix-initial vowels) is observed in stems involving extrametrical echo vowels. Again, different hiatus resolution patterns are evidenced as prominence surfaces in word-final position: in (10a-b) the underlying /u/ undergoes glide formation; in (10c) the underlying /u/ undergoes deletion; and in (10d) the underlying /u/ undergoes coalescence with the following /i/.

(10) /(CV)CVCVCV-CV/ \[CV CV CV \]
   a. pakiturwá /pakituru-a/ ‘promise (AV.IMP)
   b. tumatuturwá /tumatuturu-a/ ‘tell (AV.IMP)
   c. vantuki /vantuku-in/ ‘money (3SG.POSS)’
   d. vutukurén /vutukur-en / ‘fish (3SG.POSS)’

In addition to the type of echo vowels discussed in H. Chen (2016) and examined above, there is another type of echo vowel in Kanakanavu—the epenthetic echo vowels—which should be distinguished from the extrametrical echo vowels both in (i) how they interact differently with word prominence, and in (ii) the different roles they play when suffixation is involved. In surface forms, the epenthetic echo vowels are distinguished from the extrametrical echo vowels in showing an invariable penultimate prominence pattern. The regular penultimate prominence is observed when the words that contain them are in unsuffixed forms:

(11) (C)VCVCV(N)CV
   a. poʔotsipi ‘cook’
   b. matisáʔi ‘catch (AV)’
   c. arapiŋtsi ‘break (AV)’

When suffixation is involved, the same penultimate prominence pattern is still observed. This is the case both when the suffix forms an open syllable, as in (12), and when it forms a closed syllable, as in (13).

(12) CVCVCV
   a. poʔotsipa ‘cook (AV.IMP)’
   b. matisáʔa ‘catch (AV.IMP)’
   c. arapiŋtse ‘break (PV.NIND.NEG)’

(13) CVCVCV
   a. poʔotsipin ‘cook’
   b. matisáʔinin ‘catch (PV)’
   c. arapiŋtsin ‘break (PV)’

This invariable pattern indicates that, morphologically speaking the epenthetic echo vowels exhibit a different status from that of the extrametrical echo vowels. Whereas the latter are present in the underlying representation of the words that contain them, as represented in (9) above, the former are somehow absent in the underlying representation of words, but epenthetic only to unsuffixed forms. This is supported by the lack

---

15 The suffix -e is attached when the verb is in nonindicative negative patient/undergoer voice form (Zeitoun & Teng 2016:138).
of prominence shift observed in the suffixed forms, where no signs of hiatus resolution are observed. In this regard, the invariable penultimate prominence pattern must be attributed to the prominence-bearing vowel consistently surfacing from the underlying vowel that carries the penultimate mora. In unsuffixed forms, the echo vowel is epenthetic, serving as a separate unit following the stem:

(14) /CVCVCV(N)C-V(epenthetic)/ → [CVCVCV]<N>CV]

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In suffixed forms, on the other hand, the echo vowel is absent underlingly, and the stem ends with a consonant immediately followed by the suffix. Here, the final vowel in the stem carries the penultimate mora, hence the invariable penultimate prominence observed in the surface:

(15) /(C)VCVCV(N)C-V/ → [CVCVCV]<N>CV]

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.3 Word-final vowels and prominence positions in suffixed words

As has been demonstrated above, the different underlying statuses of word-final vowels show complex interactions with suffixation in leading to different word prominence positions, which is closely tied to whether underlying heavy syllables are formed in suffixed forms. This is summarized as in Table 3 below, in which four underlying statuses of word-final vowels are distinguished and two phonological rules are involved: Post-nasal high vowel deletion and hiatus resolution:

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Importantly, although two prominence realization patterns are observed across four categories, identical surface prominence positions may result from distinct processes. Whereas the final prominence position observed in words involving regular stem-final vowels and extrametrical echo vowels can be attributed to the formation of underlying final heavy syllables in suffixed forms, the penultimate prominence position observed in words involving final post-nasal high vowels and epenthetic echo vowels results from different phonological processes. In the former, Post-nasal high vowel deletion ensures that the underlyingly stem-final vowels are

16 Specifically, if these echo vowels were present in the underlying forms, regardless of whether they are specified as extrametrical or not, the non-existent suffixed forms *poʔotsipá or *poʔotsipjá /poʔotsipi-a/ would be predicted based on the patterns observed in (5), (6) and (10) above.

17 An alternative analysis may be to posit that the echo vowels are dropped before suffixation.
deleted in both unsuffixed forms and in the stems of the suffixed forms; in the latter, the epenthetic echo vowels are posited to be absent in the stem when suffixation is involved but inserted to the end of the stem in unsuffixed forms.

3 Prominence-attracting/-repelling morphology and prominence positions in Kanakanavu

The main goal of this section is to demonstrate that there are morphological processes in Kanakanavu that involve morphemes specified to be prominence-attracting or prominence-repelling, and that these morphemes serve as factors additional to those examined in Section 2 above that may influence prominence realization in a morphologically complex word. Specifically, it will be shown that in these processes it is solely the morphological properties of the morphemes involved, and not phonological factors, that lead to prominence positions deviating from the regular penultimate position in a word.

3.1 Prominence-attracting morphology and final prominence

Two inflectional morphemes in Kanakanavu can be identified as being prominence-attracting, namely the actor voice mild imperative suffix -an, and the patient voice mild imperative suffix -on. The two are considered as being specified to attract prominence because prominence is found to be invariably realized on the final syllable when they are suffixed onto a stem, regardless of whether there are underlying final heavy syllables resulting from the affixation or not.

Before considering -an and -on, examination of their regular/unmarked -imperative counterparts—the actor voice imperative suffix -a and the patient voice imperative suffix -o—will help recapitulate how suffixation interacts with different types of underlying word/stem-final vowels. As can be seen in (17a-c) below, final prominence is observed when -a is suffixed onto stems ending in open syllables, with the resulting final syllables in suffixed forms undergoing different hiatus resolution patterns.

(17) Final prominence from suffixation of -a onto stems ending in open syllables

- a. puʔá /puʔ-a/ ‘buy (AV.IMP)’ cf. púʔa ‘buy’
- b. pwirijá /pwirí-a/ ‘replace (AV.IMP)’ cf. pwiri ‘replace’
- c. pakiturwá /pakitur-a/ ‘promise (AV.IMP)’ cf. pakitur/pakur<u> ‘promise’

In contrast, penultimate prominence is observed in (18a-b). Here, -a is attached to stems ending in closed syllables, with (18a) showing a case of underlying final post-nasal high vowel that deletes in suffixation, and (18b) showing a case of epenthetic echo vowel that is underlyingly absent in the stem.

(18) Penultimate prominence from suffixation of -a onto stems ending in closed syllables

- a. pusuʔáná /pusuʔ-an-a/ ‘put (AV.IMP)’ cf. pusuʔán /pusuʔan/ ‘put’
- b. poʔotsípa /poʔotsip-a/ ‘cook (AV.IMP)’ cf. poʔotsípi /poʔotsip-i/ ‘cook’

The same pattern is predicted in the suffix -o, which is the patient voice counterpart of -a and may be affixed onto stems ending with different syllable types. This can be seen in (19) below.

---

18 The mild imperative adds the meaning of ‘(please) try V-ing’ to the verb it is attached to (cf. the category labeled “Directive” in the paradigm proposed by Zeitoun & Teng 2016:171). The two can be considered to serve as a polite alternative to the regular/unmarked imperative, which simply expresses commands.

19 The morpheme is observed to have two free-varying forms -o and -ow. Since the two show identical prominence realization pattern, only the former is represented in (19) due to space limitations. It is possible that -o/-ow came from Proto-Austronesian *-aw.

20 As pointed out by one of the anonymous reviewers, the underlying form of this suffix in Kanakanavu cannot be analyzed as having VV structure (i.e. the back glide [w] coming from underlying /u/) under the present framework, because it would predict invariable final prominence in (19).
(19) Final and penultimate prominence from suffixation of -o

| a. puʔó | /puʔa-o/ | ‘buy (PV.IMP)’ | cf. púʔa | ‘buy’ |
| b. pwiɾjó | /pwiɾi-o/ | ‘replace (PV.IMP)’ | cf. pwíɾi | ‘replace’ |
| c. pakitúɾó | /pakituru-o/ | ‘promise (PV.IMP)’ | cf. pakituru/pakituɾ<o>/ | ‘promise’ |
| d. pusuʔánó | /pusuʔan-o/ | ‘put (PV.IMP)’ | cf. pusuʔán/pusuʔaní/ | ‘put’ |
| e. poʔotsípo | /poʔotsip-o/ | ‘cook (PV.IMP)’ | cf. poʔotsípi /poʔotsip-i/ | ‘cook’ |

As can be seen above, in morphological processes that involve the two suffixes the prominence pattern can be fully predicted if underlying phonological representations of the stems are taken into consideration. This, however, is not the case for -an and -on, which always attract prominence to the final syllable regardless of the underlying representation of the stem. As can be seen in (20) below, prominence always lands on the final syllable when -an is attached to the stem. This is the case for both forms that end with an open syllable (20a-c), and those that end with a closed syllable (20d-e).

(20) Invariable final prominence from suffixation of -an

| a. puʔán | ‘buy (AV.MIMP)’ |
| b. pwiɾján | ‘replace (AV.MIMP)’ |
| c. pakitúɾwán | ‘promise (AV.MIMP)’ |
| d. pusuʔánán | ‘put (AV.MIMP)’ |
| e. poʔotsípán | ‘cook (AV.MIMP)’ |

This invariable prominence pattern is also observed in -on. As can be seen in (21) below, like -an, -on is found to always involve prominence realized on the final syllable in suffixed forms, regardless of the underlying representation of the stem it attaches to.

(21) Invariable final prominence from suffixation of -on

| a. puʔón | ‘buy (PV.MIMP)’ |
| b. pwiɾjón | ‘replace (PV.MIMP)’ |
| c. pakitúɾón | ‘promise (PV.MIMP)’ |
| d. pusuʔánón | ‘put (PV.MIMP)’ |
| e. poʔotsípón | ‘cook (PV.MIMP)’ |

As no evidence can be obtained to show how the invariable final prominence pattern across stems with different syllable structures may be motivated by phonological factors or processes, the pattern can only be attributed to the two mild imperative suffixes being specified as prominence-attracting.22 (20) and (21) can then be represented as in (20’) and (21’) below, where prominence is consistently realized on the two prominence-attracting suffixes,23 regardless of whether hiatus resolution is involved or not:

---

21 Notice that for pakitúɾó ‘promise (PV.IMP)’, the stem-final vowel may be analyzed as undergoing either deletion or coalescence, but no signs of glide formation is available. Compare the case when -a is involved, as in (17c).

22 As pointed out by one of the anonymous reviewers, an alternative analysis of these two suffixes could resort to a representational solution, whereby the two suffixes are simply represented with an additional (floating) mora or with a long vowel. This alternative analysis is tentatively rejected in this paper mainly because the only intra-linguistic evidence for underlying long/bimoraic vowels within a morpheme appears to be the two loan words from Taiwanese reported in H. Chen (2016), as shown in (1e-f) above. As native Kanakanavu words that have been robustly established as involving underlying final bimoraic vowels have all been attributed to historical diphthongs (/ai/, /ia/, /au/, /ua/) or a three-vowel sequence (/iau/) (H. Chen 2016:71-82), the current analysis adopts a morphological solution before definitive phonological evidence for a representational solution becomes available.

23 This is represented with the underline underneath each suffix in the underlying representation.
(20’) Invariable final prominence from suffixification of -an as a prominence-attracting suffix

a.  puʔán  /puʔa-an/  ‘buy (AV.MIMP)’
b.  pwirján  /pwiri-an/  ‘replace (AV.MIMP)’
c.  pakiturwán  /pakituru-an/  ‘promise (AV.MIMP)’
d.  pusuʔanán  /pusuʔan-an/  ‘put (AV.MIMP)’
e.  poʔotsipán  /poʔotsip-an/  ‘cook (AV.MIMP)’

(21’) Invariable final prominence from suffixification of -on as a prominence-attracting suffix

a.  puʔón  /puʔa-on/  ‘buy (PV.MIMP)’
b.  pwirjón  /pwiri-on/  ‘replace (PV.MIMP)’
c.  pakituɾón  /pakituru-on/  ‘promise (PV.MIMP)’
d.  pusuʔanón  /pusuʔan-on/  ‘put (PV.MIMP)’
e.  poʔotsipón  /poʔotsip-on/  ‘cook (PV.MIMP)’

3.2 Prominence-repelling morphology and antepenultimate prominence

There is one morphological process in Kanakanavu that can be identified as involving morphemes specified as prominence-repelling—rightward CVCV reduplication. Different from prominence-attracting morphemes, which are inflectional in nature as has been shown above, rightward CVCV reduplication is derivational, hence the lower degree of productivity in its distribution and higher degree of idiosyncrasy in its function. Despite this, all rightward CVCV reduplicants identified in Kanakanavu so far are found to occupy the penultimate and final syllables of the word after reduplication, with prominence invariably realized on the antepenultimate syllable. Functionally speaking, rightward CVCV reduplication in Kanakanavu does not itself serve as a unified category but is subsumed under several derivational processes. In the discussions that follow, one such process will be examined—in incorporated-noun reduplication.

In incorporated-noun reduplication, two morphological units are involved: (i) a verbalizer that denotes the type of event that is being expressed (e.g. po- ‘utter’, ka- ‘make’, te- ‘look for’), and (ii) a reduplicated noun that denotes an undergoer participant involved in the event. Depending on the noun stem involved, one of the three types of reduplication may obtain: (i) leftward CV reduplication (CV~), (ii) leftward CVCV reduplication (CVCV~), or (iii) rightward CVCV reduplication (~CVCV).

In the two types of leftward reduplication, as can be seen in (22) below, the reduplicant is not found to occupy either the penultimate or the final syllable. Here, the position of prominence is predicted according to the underlying syllable structure of the resulting complex word, which is found in the same position as the one in the noun stem.

(22) Leftward CV/CVCV reduplication and prominence patterns

a.  poʔɨʔîna  / po-ʔɨʔîna/  ‘tell stories’  cf. ʔîna  ‘story’
b.  pomumwáɾɨ / po-mu~mua ɾî/ ‘speak casually’  cf. mwáɾi  ‘toy’
c.  kapapaʔitsi / ka-pa~paʔitsi/ ‘make alcohol’  cf. paʔitsi  ‘alcohol’
d.  kanupanupásɨ / ka-nupa~nupasɨ/  ‘make money’  cf. nupásɨ  ‘(paper) money’
e.  kavanavanáj  / ka-vana~vanai/  ‘make/cause trouble’  cf. vanáj  ‘reason’
f.  teʔaraʔarám  / te-ʔara~ʔarami/  ‘look for meat’  cf. ʔarám  ‘meat’

In contrast, rightward CVCV reduplication involves the reduplicant ending up occupying the final and penultimate syllables of the resulting complex word, where prominence invariably falls on the antepenult. So far, three nouns have been found to involve rightward reduplication in this process: káɾi ‘word/speech’, mánu ‘child’, and ʔúma ‘farmland’. In these cases, prominence is always found on the antepenultimate syllable.

---

24 The incorporated undergoer participants are all interpreted as non-referential or indefinite (cf. Mithun 1984), and may contribute to the verbal meaning by adding a non-referential/indefinite patient, as in (22a), (22 c-e) and (23a-c), or serve a manner-modification function, as in (22b).
(23) Rightward CVCV reduplication and antepenultimate prominence

a. pokari-kari /po-kari~kari/ ‘speak’ cf. kari ‘words/speech’

b. kamanú-manu /ka-manu–manu/ ‘give birth’ cf. manú ‘child’

c. kaʔumóʔuma /ka-ʔuma~ʔuma/ ‘do farm-related work’ cf. ʔuma ‘farmland’

Note that the invariable antepenultimate prominence position involving rightward reduplication in (23) above does not appear to have any phonological motivation. On the one hand, the nominal stems themselves are prominence-bearing when occurring as independent nouns, where prominence positions are found to be predicted based on their underlying /CVCV/ syllable structure. On the other hand, there is no evidence indicating that the reduplication process affects the syllable structure of either the stem or the reduplicant which would result in antepenultimate prominence. In this regard, the pattern is attributed to the rightward reduplicant being specified as prominence-repelling, leading the prominence to shift from the originally expected penultimate syllable to the antepenultimate syllable. (23) can then be re-represented as (23’) below, where the rightward reduplicant is underlyingly specified (and represented by square brackets) as prominence-repelling:

(23’) Rightward CVCV reduplication with reduplicant represented as prominence-repelling

a. po-kaɾi~kaɾi /po-kaɾi~[kaɾi]/ ‘speak’ cf. kari ‘words/speech’

b. ka-manú–manu /ka-manu–[manu]/ ‘give birth’ cf. manú ‘child’

c. kaʔumóʔuma /ka-ʔuma~[ʔuma]/ ‘do farm-related work’ cf. ʔuma ‘farmland’

25 This word literally means “to make children”.
26 The /a/ in the noun stem becomes surrounded by /u/ after the derivation. According to Tsuchida (2003:7) it is expected to be realized as [o] in this context.
27 Accordingly, penultimate prominence should be predicted in both the original (CVCV) and reduplicated forms (expected to be CV-CVCV~CVCV). The latter, however, is not found to be the case.
28 Note here that the leftward shift appears to be the only option since there is no docking site following the prominence-repelling reduplicant for the shifted prominence to land on.
29 As pointed out by one of the anonymous reviewers, an alternative analysis would be to treat rightward CVCV reduplicants as containing a final vowel that is extrametrical, because the antepenultimate prominence would be predicted because the underlying antepenultimate syllable of a noun-incorporated verb would carry the penultimate mora (i.e. given /~ kar<i>/, the structure /po-kari–kar<i>/ would be predicted). This alternative analysis is tentatively rejected in this paper for two reasons. For one, extrametricality has only been established to be a property of root-final vowels in Kanakanavu so far, as discussed in Section 1. For the other, whereas the extrametrical status of some final vowels in Kanakanavu can be established as having a historical basis, this does not appear to be the case for the reduplicants involved in incorporated-noun reduplication. According to Tsuchida (1976:206), Kanakanavu final echo vowels (or what he refers to as “supporting vowels”) can be traced back to final vowels inserted to Proto-Tsouic words inherited from reconstructed forms that ended with final closed syllables. Many examples provided by Tsuchida clearly involve these words ending up showing antepenultimate “stress” in Kanakanavu: Proto-Southern-Formosan *təRbəs > tɨr̥əvɨsɨ ‘zelkova tree’ (ibid. 208), Proto-Hesperonesian *qpRəSc > ma-ʔirici (ibid. 207), Proto-Hesperonesian *LiməCaq > nimicəʔ ‘paddy leech’ (ibid. 208), etc. Although Tsuchida also provides Kanakanavu forms with antepenultimate prominence that resemble the reduplication pattern seen here, these do not appear to have undergone the same type of reduplication. As can be seen in the following data, the seemingly reduplicated words discussed by Tsuchida reflect reconstructed forms that appear to have already been fossilized monomorphemic words despite their seemingly reduplicative forms (i.e. They appear to have been what Zeitoun & Wu 2006:100 refer to as “lexicalized reduplication”): Proto-Hesperonesian *kiʔokə > k-um-a-kisikisi ‘to shave fur’ (ibid. 207), Proto-Austronesian *DaθDaŋ > c-um-a-caŋcəni ‘to dry by fire’ (ibid. 208). Since specification with regard to prominence-sensitivity is at least seen in two other morphemes synchronically in Kanakanavu, as examined in Section 3.1, the present paper has refrained from adopting an approach that posits extrametricality in the rightward CVCV reduplicants.
Again, this would be in sharp contrast to cases involving leftward reduplication, as has been shown in (22), where prominence position is clearly dependent on the phonology, and no morphological factors are at play in influencing prominence realization.\textsuperscript{30}

### 3.3 Morphological complexity and prominence-(in)sensitivity in prominence-repelling morphemes

As has been mentioned above, incorporated-noun reduplication in Kanakanavu is derivational in nature. It is therefore expected that when derivation of a complex word involves both rightward reduplication and some other inflectional morpheme, the former will be closer to the stem than the latter (Bybee 1985), meaning that rightward reduplication will always apply earlier to the stem than any inflectional morpheme. Interestingly, when this occurs the inflectional morpheme that attaches later will become the determinant for prominence position in the derived word, which may lead the rightward reduplicant to become prominence bearing despite its prominence-repelling status.

Consider, for example, the verb *pokarikari* ‘speak’, which becomes *pokarikarjá* ‘speak (AV.IMP)’ when attached with actor voice imperative -\textit{a}. As shown in (17a-c) above and repeated in (24a-c) below, the actor voice imperative morpheme -\textit{a} shows final prominence when it forms underlying heavy syllables with the stem-final vowel, surfacing in word-final position through hiatus resolution. This is the case even when the stem contains a prominence-repelling morpheme, such as the rightward reduplicant ~[\textit{ka}]\textit{ɾi} in (24d). Here, prominence still falls on the final syllable, which is occupied by the otherwise prominence-repelling reduplicant ~[\textit{ka}]\textit{ɾi}, with the /i/ carrying the penultimate mora in the underlying representation.

(24) Final prominence from suffixation of -\textit{a} onto stems ending in open syllables

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>puʔá /puʔa-a/</td>
<td>‘buy (AV.IMP)’ cf. puʔá ‘buy’</td>
</tr>
<tr>
<td>b.</td>
<td>pwirjá /pwirí-a/</td>
<td>‘replace (AV.IMP)’ cf. pwirí ‘replacement’</td>
</tr>
<tr>
<td>c.</td>
<td>pakituruwá /pakíturú-a/</td>
<td>‘promise (AV.IMP)’ cf. pakíturú /pakítur\textless{u}\textgreater/ ‘promise’</td>
</tr>
<tr>
<td>d.</td>
<td>pokarikarjá /pokarikari-a/</td>
<td>‘speak (AV.IMP)’ cf. pokarikari /po-kari~[kari]/ ‘speak’</td>
</tr>
</tbody>
</table>

Note that this does not mean *kari* ‘words/speech’ as a (free/bound) morpheme changes its status depending on the morphological context. Instead, its status as a morpheme remains relevant in each derivational stage but becomes irrelevant, or opaque, to the next derivational stage. In other words, as the morphological complexity increases for a stem, prominence position is determined only by the morpheme that is attached at each derivational stage.\textsuperscript{31} This accounts for the prominence alternation patterns in words sharing the morpheme *kari* ‘words/speech’ as in (25), where fully derived words with different morphological complexity and structure may involve different prominence positions.

(25) Prominence alternation patterns in words sharing *kari* ‘words/speech’

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>kári /kari/</td>
<td>‘words/speech’ Penultimate</td>
</tr>
<tr>
<td>b.</td>
<td>pokarikari /po-kari~[kari]/</td>
<td>‘speak’ Antepenultimate</td>
</tr>
<tr>
<td>c.</td>
<td>pokarikarjá /pokarikari-a/</td>
<td>‘speak (AV.IMP)’ Final</td>
</tr>
<tr>
<td>d.</td>
<td>pokarikarján /pokarikari-an/</td>
<td>‘speak (AV.MIMP)’ Final</td>
</tr>
</tbody>
</table>

### 4 Morphologically insensitive prominence in Kanakanavu

While the previous two sections have been devoted to the various interactions that prominence has with word phonology and/or morphology, the patterns discussed above have been attributed to a highly productive type of prominence that is widely observed in the Kanakanavu lexicon. In this section, more data will be presented to introduce an additional type of prominence in the language, which is not as productive as, but should arguably be distinguished from, the type of prominence that has been examined in the literature. In the

\textsuperscript{30} It is also possible that all reduplicants in Kanakanavu are prominence-repelling, and that leftward reduplicants do not lead to prominence shift because they simply happen to never occupy the final and penultimate syllables in the derived word. More systematic investigation of reduplication in Kanakanavu, however, is required to confirm this.

\textsuperscript{31} This property is also seen in how the extrametrical status of underlying echo vowels becomes irrelevant to prominence realization in suffixed forms, as shown in (10) above.
discussion that follows, the type of prominence to be introduced will be referred to as HL-prominence, which will be shown to contrast with the much more productive type of prominence discussed above (to be referred to as H-prominence) in both phonetic realization and morpho-phonological behaviors.

4.1 HL-prominence: phonetic realization and contrast with H-prominence

The label for this additional type of prominence—HL prominence—is motivated by one salient phonetic property that differs from that of H-prominence: instead of a high F₀ target, HL-prominence is found to be realized as a falling pitch contour on the penultimate syllable. As can be seen in (26), although the token number of words that have been unequivocally identified as HL-prominence-bearing is not very high, a few examples have been found to form minimal pairs or near-minimal pairs with H-prominence-bearing words.

(26) Examples of HL-prominence-bearing words

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>manâsɨ ‘probably’</td>
<td>cf. manâsɨ ‘be certain’</td>
</tr>
<tr>
<td>b.</td>
<td>makâsɨ ‘do/say like this’</td>
<td>cf. makâswa ‘do like that’</td>
</tr>
<tr>
<td>c.</td>
<td>umâra ‘take (AV)’</td>
<td>cf. umâva ‘carry on back (AV)’</td>
</tr>
<tr>
<td>d.</td>
<td>kajsîsi ‘worship/perform ritual’</td>
<td></td>
</tr>
<tr>
<td>e.</td>
<td>mjôɾu ‘pound (millet)’</td>
<td></td>
</tr>
<tr>
<td>f.</td>
<td>tanjâɾɨ ‘sun’</td>
<td>cf. tanjâɾa ‘day’</td>
</tr>
</tbody>
</table>

As can be seen in the pitch tracks in Figure 1 and Figure 2 below, the words manâsɨ ‘probably’ and makâsɨ ‘do/say like this’ both show an overall descending pitch contour occupying the entire vowel of the second syllable—[na] and [ka].

**Figure 1. manâsɨ ‘probably’**

---

32 Whether vowel duration/length also plays a role in serving as a phonetic cue for HL-prominence requires further investigation.

33 In this paper, HL-prominence is represented with a circumflex on the prominence-bearing syllable.

34 The figures presented in this paper are generated from word-list recordings of Kanakanavu words pronounced by Mr. ’Angai Ka’angaiana (翁博學).
Words bearing H-prominence, in contrast, involve the most prominent syllable being accompanied with a high pitch target that is reached. The phonetic difference can be seen in the two H-prominence-bearing words in Figure 3 and Figure 4, the former showing contrast with manâsɨ ‘probably’ in Figure 1, and the latter with makâsi ‘do/say like this’ in Figure 2. Here, despite the similar presence of a pitch fall on the penultimate syllable, there is always a high pitch target that is reached before the falling contour occurs.

Figure 2. makâsi ‘do/say like this’

Figure 3. manâsɨ ‘be certain’
4.2 Non-interaction with underlying morphological representations

Compared to the distribution of H-prominence, that of HL-prominence is far less productive, as it appears to only occupy a small part of the Kanakanavu lexicon. So far, the words identified to bear HL-prominence, as in (26), are predominantly verbal predicates, and there appear to be no semantic motivations that may unite HL-prominence-bearing words as a class, nor any common morphological property.

Despite the idiosyncrasy in the distribution, one clear behavioral difference can be observed between HL-prominence and H-prominence, and that lies in the role that morphology plays. Whereas H-prominence has been shown to have various interactions with phonological and morphological representations of words, HL-prominence is found to be invariably realized on the penultimate syllable, and it occurs only in fully derived words that are lexically specified to be HL-prominence-bearing.

This may be illustrated by examining words that appear to be morphologically related to HL-prominence-bearing words but are nonetheless H-prominence-bearing. To begin with, HL-prominence-bearing words that are clearly morphologically simplex may serve as the stem from which other complex words are derived. In these cases, however, words in the latter category are found to be H-prominence-bearing, and the prominence position would shift according to the phonological structure of the derived word. This can be seen in (27), where the two HL-prominence-bearing words *kajsîsi* ‘perform ritual’ and *mjôru* ‘pound (millet)’ serve as stems for the two nominalized forms *takajsisjá* ‘deity/ritual place’ and *mjorwá* ‘pounded millet’, respectively. Here, the latter two words are H-prominence-bearing, and involves prominence to fall on the final syllable, which is predicted from the underlying syllable structure.

(27) Examples of H-prominence-bearing words derived from HL-prominence-bearing words

a.  takajsisjá  /ta-kajsisi-a/  ‘deity/ritual place’  cf. kajsîsi ‘to worship/perform ritual’

b.  mjorwá  /mjorut-a/  ‘pounded millet’  cf. mjôru ‘to pound (millet)’

The data in (27) above may give the impression that HL-prominence is always idiosyncratically assigned to the stem that is specified for it, and that as the morphological complexity increases, H-prominence overrides. This, however, is not always the case because there are also HL-prominence-bearing words that can be analyzed as morphologically complex but share the same morpheme with other complex words that are nevertheless H-prominence-bearing. This can be seen in (28) below, where the HL-prominence-bearing word *umâra* ‘take (AV)’ shares the stem *ara* ‘take’ with other H-prominence-bearing words. Notice here that all H-prominence-bearing words derived from the stem *ara* except for *umâra* ‘take (AV)’ involve H-prominence being assigned according to word syllable structure, just like in the cases discussed in Section 2.
(28) Prominence patterns observed in words analyzable as sharing the stem aɾa ‘take’

a. umâra /ara-in/ ‘take (AV)’
   Penultimate HL
b. arin /ara-in/ ‘take (PV)’
   Final H
c. sjára /si-ara/ ‘tool for taking / take (INS.NMLZ)’
   Penultimate H
d. ?apára /?apa-ara/ ‘make … take / take (CAUS)’
   Penultimate H

A contrast between umâra ‘take (AV)’ and umáva ‘carry on back (AV)’ [AV-carry.on.back] as in (26c) shows that a morphologically complex HL-prominence-bearing word may also be analyzed as sharing an affix (the actor voice morpheme um in this case) with another word, but again, HL-prominence is only found on the former, and not the latter.

Finally, the presence of the minimal pair manâs ‘probably’ vs. manás ‘be certain’ presented in (26a) above indicates that the morphological similarity between an HL-prominence-bearing word and an H-prominence bearing word may only be a result of surface resemblance, and that the former may not be productively linked to the latter in the lexicon.\[^{35}\] This is supported by the observation that despite surface resemblance, an HL-prominence-bearing word may differ from an H-prominence-bearing word in word class, and that H-prominence-bearing words may share not only morphological units, but also word class. Although manâs ‘probably’ may be analyzed as showing the same structure as manás ‘be certain’, the former has been shown to serve as a propositional modal predicate (Cheng & Sung 2015:34-35), which shows systematic behavioral differences from verbs/proposition-level predicates (ibid. 40). In contrast, the H-prominence-bearing manás ‘be certain’ is clearly productively linked to the other H-prominence-bearing words that serve as verbs and share the stem -nasi ‘certain’:

(29) H-prominence-bearing words sharing the stem -nasi ‘certain’

a. manás ‘be certain / certain (AV.STAT)’ /ma-nasi/
b. munás ‘move directly / move-certain’ /mu-nasi/
c. ponás ‘speak directly/ utter-certain’ / po-nasi/

In this regard, HL-prominence can be concluded as a type of prominence that is invariably realized on the penultimate syllable of words that are specified to be HL-prominence-bearing. Crucially, HL-prominence-bearing words are listed as single lexical entries in the lexicon. Despite sharing formal resemblances on the surface, they are only productively linked to H-prominence-bearing words when they serve strictly as the stem to which affixes are attached to derive morphologically complex words that are H-prominence-bearing.\[^{36}\] The words in (26) can then be represented as in (26’):

(26’) Posited underlying representations of HL-bearing-words

a. manâs /manâs/ ‘probably’
b. makâsi /makâsi/ ‘do/say like this’
c. umâra /umâra/ ‘take’
d. kajsi /kajsi/ ‘perform ritual’
e. mjôru /mjôru/ ‘pound (millet)’
f. tanjâri /tanjâri/ ‘sun’

5 Word-level prominence in Kanakanavu: cyclic and postcyclic processes

In this section, Kanakanavu prominence patterns observed in the literature will be integrated with the data discussed in the previous three sections of the present paper, whereby a revised analysis of Kanakanavu word-level prominence couched in Lexical Phonology (Kiparsky 1982, Mohanan 1982, Booij & Rubach 1987) will

\[^{35}\] I thank Dr. Matthew Gordon for pointing this out to me.

\[^{36}\] It is likely that productive linkages between contemporary HL-prominence-bearing-words and some H-prominence-bearing words existed historically (and that the former used to be H-prominence-bearing as well), but was eliminated by a process whereby morphologically complex forms were reanalyzed as monomorphemic forms, accompanied by the rise of word-level HL-prominence.
be proposed. Specifically, the observed prominence patterns will be attributed to two types of tones—(i) the **H tone** that is assigned as a result of complex interactions of cyclic and postcyclic rules in the derivation of morphological simplex and complex forms, and (ii) the **HL tone** that is underlyingly specified in morphologically simplex forms that may or may not be productively linked to other H-prominence-bearing forms.

The majority of the discussions that follow will be dedicated to demonstrating that the varying positions of H-prominence, which has been shown to be sensitive to word phonology and morphology, are attributed to three H tone-related rules that apply cyclically in interacting with other cyclic and postcyclic rules in the derivation. The invariable penultimate pattern observed for HL-prominence, on the other hand, is attributed to an underlying HL-tone that may or may not block assignment of the H tone depending on whether suffixation is involved.

### 5.1 H-prominence and the interactions between cyclic and postcyclic rules in Lexical Phonology

To capture the complex interactions between H-prominence and the various factors discussed in the previous sections, the present analysis posits a rule **H-assignment**, which assigns a H tone to the penultimate mora in the input form. As indicated by the different prominence patterns observed in the morphologically simplex words discussed in H. Chen (2016) (reviewed in Section 1), **H-assignment** interacts with at least two other phonological rules—**Post-nasal high vowel deletion** and **Hiatus resolution**—whose application must follow that of **H-assignment**. The ordering of the three rules can be illustrated as in (30), where **H-assignment** assigns an H tone to (i) /a/ in the input /punai/ ‘pigeon’, the second /a/ in the input /tsanaa/ ‘field (wet rice)’ and the first /i/ in the input /makasini/ ‘now’ due to their underlyingly penultimate positions, and to (ii) the second /u/ in the input /vutukur<u>/ ‘fish’ due to the underlying extrametrical status of the final /<u>/.

**Crucially, Post-nasal high vowel deletion** deletes the final /i/ in /makasini/ after **H-assignment** applies, leading to the H tone surfacing in final position, and **Hiatus resolution** coalesce /a/ and /i/ in /punai/ after the /a/ has been assigned a H tone, which then surfaces in final position.

(30) Derivation of **puné** ‘pigeon’, **tsaná** ‘field (wet rice)’, **vutúkuɾ</u>** ‘fish’ and **makasín** ‘now’

<table>
<thead>
<tr>
<th>Input</th>
<th>punái</th>
<th>tsanáa</th>
<th>vutukur&lt;u&gt;</th>
<th>makasini</th>
</tr>
</thead>
<tbody>
<tr>
<td>puné</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>makasín</td>
</tr>
<tr>
<td>tsaná</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Output</td>
<td>puné</td>
<td>tsaná</td>
<td>vutukuru</td>
<td>makasín</td>
</tr>
<tr>
<td>(Final H)</td>
<td>(Final H)</td>
<td>(Antepenult. H)</td>
<td>(Final H)</td>
<td></td>
</tr>
</tbody>
</table>

As implied in the H-prominence patterns shown by suffixed words involving stem-final vowels with different underlying statuses (discussed in Section 2.1), **H-assignment** and **Post-nasal high vowel deletion** are by definition **cyclic** rules because they are found to be applied in tandem with the morphological process of suffixation (Rubach 2008:462). In contrast, **Hiatus resolution** is a **postcyclic** rule because it does not interact with morphology (hence not applied cyclically) but plays a role in the surfacing of final heavy syllables of fully derived forms, either when they result from an underlying /VV/ sequence or are created through suffixation (Rubach 2008:466-467). In order to account for the prominence shift patterns observed in some suffixed forms, an additional cyclic rule—**Culminativity**—is posited to follow **H-assignment**: if a form receives more than one H tone due to the cyclic application of **H-assignment**, **Culminativity** deletes all but the rightmost H tone before the form enters the next stage of the derivation.

As can be seen in (31), the cyclic applications of **Post-nasal high vowel deletion** and **Culminativity** account for the correct surface H-prominence patterns in **pusuʔana /pusuʔan-a/’ put (AV.IMP)** and **puʔa /puʔa-a/’ buy**

---

37 *Hiatus resolution* had been posited as a postlexical rule in an earlier version of this paper. However, as pointed out by one of the anonymous reviewers, this would imply its ability to apply across word boundaries (Rubach 2008:461), which is not observed in the author’s data.

38 This is motivated by the observation that although the cyclic nature of **H-assignment** results in multiple H tones in a form in the derivation, it is always the rightmost H tone that is accountable for the surface prominence position.
Here, the underlyingly penultimate /a/ in /pusuʔan/ is assigned a H tone before the deletion of post-nasal /ɨ/ by Post-nasal high vowel deletion on the first cycle in the cyclic component; it blocks the application of H-assignment on the second cycle and surfaces as the penultimate H-prominence-bearing vowel. The stem-final /a/ in /puʔa-a/ receives a H tone that survives through the derivation due to the application of Culminativity on the second cycle; it eventually coalesces with the following /a/ due to the application of Hiatus resolution in the postcyclic component. This mechanism therefore accounts for the (lack of) H-prominence shift across unsuffixed and suffixed forms involving the stems /pusuʔán ‘put’ and /puʔa ‘buy’, as shown in (31).

(31) Derivation of /pusuʔán ‘put’, /pusuʔána ‘put (AV.IMP)’, /puʔa ‘put’ and /puʔá ‘put (AV.IMP)’

**Cyclic Component**

<table>
<thead>
<tr>
<th>Cycle 1</th>
<th>pusuʔání</th>
<th>pusuʔání</th>
<th>puʔa</th>
<th>puʔa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle 2</td>
<td>pusuʔán-a</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Postcyclic Component**

<table>
<thead>
<tr>
<th>Output</th>
<th>pusuʔán</th>
<th>pusuʔána</th>
<th>puʔa</th>
<th>puʔá</th>
</tr>
</thead>
</table>

5.2 The (lack of) effect of underlying morpheme statuses

One further step of elaboration of the current model concerns the observed limitation of underlying statuses of morphemes to each stage of the derivation, as has been seen in the loss of extrametricality effect from underlyingly stem-internal echo vowels in (10), and that of prominence-repelling status of rightward CVCV reduplicants in (23). The configuration as illustrated in (31) should therefore be further stipulated such that any morphological information inherited from the input is eliminated once a form enters the next stage of the derivation, where morphological information includes morpheme boundaries and underlying morpheme specifications.

The stipulation allows the model to capture the prominence shift patterns observed across unsuffixed and suffixed forms involving echo vowels (examined in Section 2.2). To begin with, the different statuses of echo vowels are captured by (i) positing the underlyingly stem-internal echo vowels to be present but specified as extrametrical in the input to the first cycle of the derivation, and (ii) treating the stem-external echo vowels as not involved until the second cycle of the derivation is reached. As can be seen in (32), the correct forms /pakítuɾu ‘promise’, poʔotsípi ‘cook’ and poʔotsípa ‘cook (AV.IMP)’ are derived with the former undergoing only one cycle of the derivation in the cyclic component but the latter two going through both cycles due to the involvement of the epenthetic vowel -i and the suffix -a. In addition, the extrametricality of underlying /<u>/ in the suffixed form /pakituswa ‘promise (AV.IMP)’ is ensured to only cast its influence on the first cycle, as the model prevents it from being carried over onto the second cycle. In this regard, the stem-final /u/ manages to get assigned a H tone by H-assignment on the second cycle, leading to H-prominence surfacing in the final syllable after Hiatus resolution applies in the postcyclic component.

---

39 Post-nasal high vowel deletion is not included in (32) and other derivations below due to space limitation and its consistent lack of application in the derivation of the forms discussed.
(32) Derivation of *pakítuɾu* ‘promise’, *pakítuɾwá* ‘promise (AV.IMP), *poʔotsípi* ‘cook’ and *poʔotsípa* ‘cook (AV.IMP)’

**Cyclic Component**

<table>
<thead>
<tr>
<th>Cycle 1</th>
<th>pakítuɾ&lt;sub&gt;u&lt;/sub&gt;</th>
<th>pakítuɾ&lt;sub&gt;u&lt;/sub&gt;</th>
<th>poʔotsíp</th>
<th>poʔotsíp</th>
<th>H-assignment</th>
<th>Culminativity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle 2</td>
<td>pakítuɾu-a</td>
<td>poʔotsíp-i</td>
<td>poʔotsíp-a</td>
<td></td>
<td>H-assignment</td>
<td>Culminativity</td>
</tr>
</tbody>
</table>

**Postcyclic Component**

<table>
<thead>
<tr>
<th>pakítuɾu</th>
<th>pakítuɾwá</th>
<th>poʔotsípi</th>
<th>poʔotsípa</th>
</tr>
</thead>
</table>

**Output**

<table>
<thead>
<tr>
<th>pakítuɾu</th>
<th>pakítuɾwá</th>
<th>poʔotsípi</th>
<th>poʔotsípa</th>
</tr>
</thead>
</table>

(33) Derivation of *poʔotsípa* ‘cook (AV.IMP)’, *poʔotsípán* ‘cook (AV.MIMP)’, *pwíɾjá* ‘return (AV.IMP)’ and *pwíɾján* ‘return (AV.MIMP)’

**Cyclic Component**

<table>
<thead>
<tr>
<th>Cycle 1</th>
<th>poʔotsíp</th>
<th>poʔotsíp</th>
<th>pwíɾjí</th>
<th>pwíɾjí</th>
<th>H-assignment</th>
<th>Culminativity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle 2</td>
<td>poʔotsíp-a</td>
<td>poʔotsíp-an</td>
<td>pwíɾjí-a</td>
<td>pwíɾjí-an</td>
<td>H-assignment</td>
<td>Culminativity</td>
</tr>
</tbody>
</table>

**Postcyclic Component**

<table>
<thead>
<tr>
<th>poʔotsípa</th>
<th>poʔotsípán</th>
<th>pwíɾjí</th>
<th>pwíɾjíán</th>
</tr>
</thead>
</table>

**Output**

<table>
<thead>
<tr>
<th>poʔotsípa</th>
<th>poʔotsípán</th>
<th>pwíɾjí</th>
<th>pwíɾjíán</th>
</tr>
</thead>
</table>

5.3 Prominence-attractingness and prominence-repellingness

To capture the H-prominence patterns observed in words containing morphemes with prominence-sensitive specifications—i.e. morphemes that are specified as prominence-attracting or prominence-repelling (investigated in Section 3)—two more cyclic rules are further posited to apply after H-assignment. First, the rule *H-attraction* applies when a prominence-attracting morpheme is present in the input, whereby a H tone assigned by H-assignment on the current cycle gets delinked from the originally assigned mora and relinked to the mora carried by the prominence-attracting morpheme. As can be seen in (33) below, the invariable final prominence in words suffixed with the prominence-attracting -an such as *poʔotsípán* ‘cook (AV.MIMP)’ and *pwíɾján* ‘replace (AV.MIMP)’ is attributed to the application of H-attraction on the second cycle, which leads the assigned H tone to always surface on the final syllable, regardless of whether Hiatus resolution applies in the postlexical component or not. This contrasts with the derivation of words suffixed with the non-prominence-attracting -a, as in *poʔotsípa* ‘cook (AV.IMP)’ and *pwíɾjá* ‘return (AV.IMP)’, in which the surface positions of H-prominence are in large part attributed to the application of H-assignment on the second cycle.

---

40 *H-attraction* can be implemented as an autosegmental spreading-cum-delinking rule, whereby a H tone spreads rightward to a mora that includes a [+Exception] feature and delinks with the original mora to which it had been linked, where there are no linked tones intervening between the two moras.
Second, the rule H-repellence applies when a H tone assigned on the current cycle is linked to a mora carried by a prominence-repelling morpheme, whereby the H tone gets delinked from that mora and relinked to the closest mora that is not carried by any prominence-repelling morpheme.\textsuperscript{41} In (34), it can be seen that the derivation of the correct form pokarikari ‘speak’ depends on the application of H-repellence on the second cycle, which is where a H tone gets assigned by H-assignment onto the /a/ vowel of the prominence-repelling reduplicant ~[kari]. Here, H-repellence relinks the H tone to the preceding vowel /i/ located outside the reduplicant, leading the tone to eventually surface on the antepenultimate syllable. Notice, however, that the prominence-repelling status of the reduplicant loses its influence on H-prominence position once a form containing it enters the third cycle. In the derivation of morphologically more complex forms such as pokarikaryá ‘speak (AV.IMP)’ and pokarikarián ‘speak (AV.MIMP)’, therefore, the surface prominence positions are attributed to H-assignment and/or H-attraction applying on the third cycle, where Culminativity ensures that it is the newly assigned H tone that survives, and not the one relocated to the first /i/ on the second cycle.

\begin{table}[h]
\centering
\begin{tabular}{lllll}
\hline
& Cycle 1 & Cycle 2 & Cycle 3 & Postcyclic Component \\
\hline
\textbf{Cyclic Component} & & & & \\
\kari{kari} & \kari{kari} & \kari{kari} & \kari{kari} & H-assignment \\
\kári{kári} & \kári{kári} & \kári{kári} & \kári{kári} & H-attraction \\
\hline
\textbf{Postcyclic Component} & & & & \\
\kari{kari} & pokarikari & pokarikari-a & pokarikari-an & Hiatus resolution \\
\hline
\textbf{Output} & & & & \\
\textit{(Penult. H)} & pokarikari & pokarikari-já & pokarikari-ján & \\
\textit{(Antepen. H)} & & & & \\
\textit{(Final H)} & & & & \\
\textit{(Final H)} & & & & \\
\hline
\end{tabular}
\end{table}

5.4 HL-prominence from underlyingly specified HL-tone

Finally, the current model treats HL-prominence as a type of prominence that is present as a HL tone in the underlying representation of HL-prominence-bearing words. This is motivated by the observation that HL-prominence is invariably realized on the penultimate syllable, and that HL-prominence-bearing words do not appear to be productively linked to H-prominence-bearing words that show surface formal resemblance, unless the former serve as stems from which the latter are derived.

As can be seen in the derivation in (35)\textsuperscript{42} below, the surface penultimate HL-prominence observed in the morphologically simplex forms kajsíši ‘perform ritual’ and mjórú ‘pound (millet)’ is attributed to (i) an underlyingly present HL tone in the input and (ii) H-assignment failing to apply on the first cycle due to the HL tone being linked to the penultimate mora in the input form. However, once a form enters the second cycle via affixation, H-assignment can apply if a final heavy syllable is created. The derivation of complex forms

\textsuperscript{41} In the same vein, H-repellence can be implemented as an autosegmental spreading-cum-delinking rule, whereby a H tone linked to a mora contained in a rightward CVCV reduplicant spreads to the nearest mora that is not contained in the reduplicant and delinks with the original mora to which it had been linked.

\textsuperscript{42} H-attraction and H-repellence are not included in (35) and (36) due to space limitation and their consistent lack of application in the derivation of the forms discussed.
such as takajsísá ‘deity/ritual place’ and mjórwá ‘pounded millet’ therefore involves H-assignment assigning a H tone on the second cycle and Culminativity deleting the underlyingly specified HL tone, with the newly assigned H tone surviving to the postcyclic component where Hiatus resolution applies before surfacing in the output.

(35) Derivation of kajsísí ‘perform ritual’, takajsísá ‘deity/ritual place’, mjóru ‘pound millet’ and mjórwá ‘pounded millet’

<table>
<thead>
<tr>
<th>Cyclic Component</th>
<th>Cycle 1</th>
<th>Cycle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>kajsísí</td>
<td>kajsísí</td>
<td>ta-kajsísí-a</td>
</tr>
<tr>
<td>mjóru</td>
<td>mjóru</td>
<td>mjórwá</td>
</tr>
<tr>
<td>H-assignment</td>
<td>Culminativity</td>
<td>H-assignment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Postcyclic Component</th>
<th>umára</th>
<th>ará-in</th>
</tr>
</thead>
</table>
| kajsísí              | takajsísá | ára-
| mjóru                | mjórwá    | -      |
| Hiatus resolution    |         | Culminativity |

<table>
<thead>
<tr>
<th>Output</th>
<th>umára</th>
<th>ará-in</th>
</tr>
</thead>
</table>
| kajsísí| takajsísá | ará-
| mjóru | mjórwá  | -      |
| (Penult. HL) | (Final H) | (Penult. HL) | (Final H) |

Under this analysis, the lack of productive linkage between a HL-prominence-bearing word and H-prominence-bearing words that show surface formal resemblance is captured with (i) the inert behavior of the HL tone reflecting its underlying penultimate position blocking H-assignment, and (ii) the cyclically assigned H tone actively interacting with word formation rules. This difference can be illustrated as in (36) below, where the HL-prominence-bearing word umára ‘take’ involves the underlyingly specified HL tone blocking all the H-tone-related cyclic rules from applying, which is in sharp contrast to H-prominence-bearing words sharing the stem ara ‘take’, where the H-prominence patterns are a result from complex interactions of different cyclic and postcyclic rules.

(36) Derivation of umára ‘take’ and H-prominence-bearing words sharing the stem ara ‘take’

<table>
<thead>
<tr>
<th>Cyclic Component</th>
<th>Cycle 1</th>
<th>Cycle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>umára</td>
<td>ara</td>
<td>ára-in</td>
</tr>
</tbody>
</table>
| ara              | ára     | ára-
| ara              | -       | -      |
| H-assignment     | Culminativity | H-assignment |

<table>
<thead>
<tr>
<th>Postcyclic Component</th>
<th>umára</th>
<th>ará-in</th>
</tr>
</thead>
<tbody>
<tr>
<td>umára</td>
<td>arán</td>
<td>si-ára</td>
</tr>
<tr>
<td>ará-in</td>
<td>ára-</td>
<td>-</td>
</tr>
</tbody>
</table>
| ¿apa-
| ¿apára  | -      |      |
| Hiatus resolution    |         | Culminativity |

<table>
<thead>
<tr>
<th>Output</th>
<th>umára</th>
<th>ará-in</th>
</tr>
</thead>
<tbody>
<tr>
<td>umára</td>
<td>arán</td>
<td>si-ára</td>
</tr>
<tr>
<td>ará-in</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
| ¿apa-
| ¿apára  | -      |      |
| (Penult. HL) | (Final H) | (Penult. H) | (Final H) |

6 Summary and conclusion
This paper has attempted to engage with an ongoing discussion in Formosan linguistics on word-level prosodic prominence in Kanakanavu. While previous studies have made contributions to our understanding of prominence in the language from phonetic and phonological perspectives, the above discussions have demonstrated that the prominence realization patterns observed from additional data are a result of neither phonological nor morphological factors alone. Instead, it is through investigating the complex interactions between phonology and morphology that have led to further insights into word-level prosodic prominence in the language.

43
In particular, it has been shown that Kanakanavu involves (i) different morpho-phonological processes giving rise to prominence realization and shift patterns and (ii) (at least) two prominence types, each exhibiting specific phonetic properties and morpho-phonological behaviors. As discussed in sections 1, 2 and 3, the surface realization patterns of H-prominence are not only concerned with the underlying representation of forms, but also influenced by affixation involving morphemes that may or may not have prominence-related specifications. In Section 4, the invariably penultimate HL-prominence is discussed with regard to its lack of interaction with morphology and the general lack of productive linkage between words that show surface formal resemblance but bear different prominence types. The differing properties of H- and HL-prominence are captured in an integrated analysis couched in Lexical Phonology in Section 5. The former is analyzed as involving a cyclic rule of \textit{H-assignment}, whose application assigns a H tone to the input form and interacts with other cyclic and postcyclic rules in the derivation. The latter, on the other hand, is analyzed as an underlyingly present HL tone that may or may not block the cyclic application of \textit{H-assignment}.

Whereas previous studies have witnessed a debate over whether to refer to Kanakanavu word-level prominence as “stress”, the discussion in this paper in no way leads to a simple classification of the language as having a prototypical “stress” or “tone” system. Instead of employing the generic label “pitch-accent”, however, the current analysis makes it possible for a property-driven approach (Hyman 2009) to be taken, where Kanakanavu can be discussed with regard to the extent to which it exhibits pitch features (“tone”) and/or metrical structure (“stress”) at the word level. On the one hand, the language does show word-level prosodic properties that require reference to pitch features. It has been shown that H-prominence is realized as a high pitch target, and that HL-prominence is realized as a falling pitch contour. Although the surface prominence-realizing unit is the syllable (the prototypical stress-bearing unit), it is clearly the mora (the prototypical tone-bearing unit) that is the \textit{de facto} prominence-bearing unit, as it has been repetitively demonstrated that identifying the underlyingly penultimate mora predicts surface realization of prominence.

On the other hand, whereas the observed patterns do not lead to an analysis that posits (H-)prominence to be assigned metrically (a prototypical property of stress systems), the distribution of the H and HL tones is apparently restricted to some extent. First, despite the scarcity of minimal pairs showing a H- vs. HL-prominence contrast, the only position where such a contrast can be found is the surface penultimate syllable. Second, under the current analysis Kanakanavu can be characterized as having word-level prominence that is both obligatory and culminative. As may have been observed throughout this paper, every content word in the language surfaces with at least one pitch-specified prominence, and this has been analyzed as resulting from either the cyclic assignment of the H tone or the underlying specification of the HL tone. In addition, every word surfaces with at most one pitch-specified prominence as well. There are no Kanakanavu words considered so far that unequivocally show more than one prosodically prominent syllable, and the proposed model crucially involves the rule Culminativity to ensure that forms are linked with only one pitch feature (H or HL) before surfacing. Finally, the current analysis also implies that Kanakanavu exhibits a \textit{privative} /H, HL, \emptyset/ tone distinction. The underlyingly penultimate mora either receives a H tone or is specified to be linked with HL a tone, with the rest of the moras in a word being left unspecified for tone.

In this regard, the present paper has hopefully made the following specific contributions. In addition to discussing the phonetic properties of prominence realization in Kanakanavu (which is the main focus in S. Chen 2016), this paper has further explored different prominence types that not only show surface phonetic differences but provide a window to distinct morpho-phonological behaviors. Moreover, it has been demonstrated that, in addition to the phonological factors at play (as examined in H. Chen 2016), word-level prosodic prominence in Kanakanavu is essentially a morpho-phonological phenomenon, the investigation of which necessarily involves understanding of the rich morpho-phonology in the language. Regardless, the proposed analysis couched in Lexical Phonology has enabled a property-driven approach in characterizing/typologizing the prominence system in Kanakanavu, which, under the current analysis, appears

---

43 This is also pointed out in S. Chen & Sung’s (2016) phonetics-based study on Kanakanavu word prominence.

44 Compare, however, H. Chen’s (2016:94-97) proposed Kanakanavu stress algorithm, which involves moraic trochees being formed iteratively from right to the left.
as a restricted tone system that shows obligatoriness and culminativity at the word-level, with a privative /H, HL, Ø/ distinction.\footnote{As pointed out by one of the anonymous reviewers, prosodic systems discussed in Zec (1999) and Köhnlein (2016) may be comparable to some degree to the Kanakanavu system discussed in this paper.}

Although more systematic investigations of Kanakanavu word-level prominence are surely required, two directions hinted/suggested by recent studies seem promising for future research. First, one of S. Chen’s (2016) findings in her acoustic study is that both pitch and vowel duration play an important role as phonetic cue for prominence in Kanakanavu. While this study has mainly focused on pitch features as the key phonetic differences between H- and HL-prominence, it still behooves for a phonetics-based study to be conducted in order to enunciate the relationships between pitch and vowel duration in the two types of prominence identified. Second, one of the major discoveries made by H. Chen (2016) is that Kanakanavu has undergone large-scaled phonological changes, mainly through the process of monophthongization. This has led her to conduct a complete reassessment of its phonological system. Historically speaking, therefore, it will be worthwhile to investigate whether the opposition of two prominence types has its origin in drastic segmental changes that the language has undergone. The next steps for a continued exploration of Kanakanavu word-level prosody as suggested here would, of course, require further systematic data collections and collaborative documentation of the language. They are therefore outside the scope of this paper and left for future research endeavors.

References


Zeitoun, Elizabeth and Chen-huei Wu. 2006. An overview of reduplication in Formosan languages. Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday, ed. by Henry Y. Chang, Lillian M. Huang, and Dah-an Ho, 97-142. Taipei: Academia Sinica.
THE DISTRIBUTION OF THE MĀORI

GENITIVE RELATIVE CONSTRUCTION¹

Michaela Socolof     Junko Shimoyama
McGill University     McGill University
michaela.socolof@mail.mcgill.ca  junko.shimoyama@mcgill.ca

Abstract
This paper focuses on the genitive relative construction (GRC) in Māori. The GRC is a type of relative clause construction that can be used to relativize non-subject DP positions, and it has been described in the literature as a means of getting around an extraction restriction on direct objects. We show that the distribution of the GRC is wider than generally described, and we argue that a split-ergative analysis of Māori, following Pucilowski (2006), better captures this distribution.

Keywords: genitive relative construction, relative clauses, genitive subjects, Polynesian, Māori, syntax

ISO 639-3 codes: mri, jpn

1 Introduction: the genitive relative construction (GRC) in Māori
Māori, an Eastern Polynesian language, allows a relative clause structure that has been called the genitive relative construction (GRC). The GRC, found in several other Polynesian languages as well, is generally characterized as follows: the agent of the relativized clause surfaces with genitive marking and appears to possess the head noun. An example is given in (1).²

(1) Ka mōhio ahau ki te tangata a Hone [i kōhuru ai ∅ subj ∅ obj]
  TAM know 1SG OBJ the man of John TAM murder PART
  ‘I knew the man that John murdered.’ (Bauer et al. 1997:570)

Genitive subjects in relative clause constructions in Altaic languages have attracted attention (see, for example, Krause 2001, Kornfilt 2008, Miyagawa 2011); an example of such a sentence in Japanese is given in (2). The GRC in Polynesian, however, is an understudied construction, and it raises interesting syntactic and semantic puzzles. Some of the first attempts to provide a detailed formal syntactic analysis of the Polynesian GRC were Herd et al. (2005), Herd et al. (2011), and Otsuka (2010), which discuss several properties of the construction.

(2) [Mary-ga/no aishiteiru] otoko-o mita. [Japanese]
  Mary-NOM/GEN love.PRS.PROG man-ACC saw.
  ‘I saw the man who Mary-NOM/GEN loves.’ (Krause 2001:36)

¹ Many thanks to Sandra Chung, Michael Erlewine, Vincent Homer, Henrison Hsieh, Maria Polinsky, and Lisa Travis, as well as the syntax reading group at McGill, the AFLA 25 audience and organizers, two anonymous AFLA abstract reviewers, and two reviewers of the paper for valuable feedback. The research reported here is supported in part by SSHRC (435-2015-0454), to which we are grateful.

² The following abbreviations are used in glosses: AGT = agent, CIA = passive/Pattern II, DIR = directional, PART = particle, PERS = pronoun/proper name, PREP = preposition, TAM = tense/aspect/mood. All other glosses follow the Leipzig glossing rules.
In this paper, we will provide a characterization of the GRC’s distribution, with a focus on two main questions: (i) What is it that gets genitive-marked? and (ii) What position gets relativized? We will present two puzzles regarding the distributional patterns of the GRC, and we will show that one of those puzzles goes away if we adopt Pucilowski’s (2006) analysis of Māori as a split-ergative language, in turn providing support for that analysis. Finally, we will show that the GRC has a wider distribution than generally assumed, and we will discuss the implications of these facts on syntactic analyses of the construction.

We leave as a topic for future research the mechanism that allows the subject of a relative clause to end up with genitive case outside of the clause. Furthermore, we restrict our focus to the GRC in Māori and its relationship with other aspects of Māori syntax, and leave comparative work on other Polynesian languages for future work as well.

2 Relative clause formation in Māori

Māori is a predominantly VSO language with multiple strategies for relativization, which have been described in detail by Bauer et al. (2003). The primary relative clause strategy, known as the ‘gap strategy’, has the relativized position realized as a gap inside the relative clause, as in (3). Schematically, we can represent this as in (4).

(3) \[ Ko tēnei te tangata [i tae tōmuri mai \empt\]. \]
PREP DEM.PROX the person TAM arrive late DIR
‘This is the man who arrived late.’ (Kelly 2015:67)

(4) \[ D NP [Op_1 [V \empt \empt (DP_{obj})]] \]

The GRC, as described by Bauer et al. (2003), differs from simple relativization in two main respects. First, the GRC has two gaps – in addition to the expected gap for the relativized position, it has a gap for the agent (or, as we will see later, the experiencer) of the relative clause verb. Second, the agent surfaces to the left of the relative clause, marked with genitive case, as in (6). The genitive-marked DP thus appears to be in a possessor-possessee relationship with the head of the relative clause. The simple, non-relativized version of (6) is provided in (5).

(5) \[ I \hohuru a Hone i te tangata. \]
TAM murder PERS John OBJ the man.
‘John murdered the man.’

(6) \[ Ka mōhio ahau ki te tangata a Hone [i kōhuru ai \empt \empt]. \]
TAM know 1SG OBJ the man of John TAM murder PART
‘I knew the man that John murdered.’ (Bauer et al. 1997:570)

It should be noted that relative clauses that use the GRC are obligatorily marked with the postverbal particle \empt\ (however, this particle is suppressed in the presence of the postverbal deictic markers \empt\, \empt\, and \empt\, as well as the imperfective marker \empt\).

The GRC in Māori may also occur without an overt head noun, as in (7), (8), and (9). In these cases, the head noun does not appear in the sentence at all.

(7) \[ Kia mea-tia tāu \empt \[ [e pai ai \empt \empt]. \]
TAM do-CIA 2SG.POSS TAM like PART
‘Thy will be done (‘Let that be done which you like.’).’ (Harlow 2007:186)
### (8) ... ko tā taku ringa Ø [i ngaki ai Øsubj

<table>
<thead>
<tr>
<th>TOP</th>
<th>the-of 1SG.POSS hand TAM cultivate PART</th>
</tr>
</thead>
<tbody>
<tr>
<td>Øobj</td>
<td>me waiho tēna ki a au. TAM leave DEM.PROX to PERS 1SG</td>
</tr>
</tbody>
</table>

‘...what my hand has cultivated, that should be left for me.’ (Bauer et al. 1997:583)

### (9) Me whakarongo tātou ki a rātou Ø [e mea nei Øsubj Øobj].

| TAM listen 1PL.INCL to the.of 3PL TAM say PROX |

‘We’d better listen to what they are going to say.’ (Bauer et al. 1997:583)

In GRC configurations, a possessor of the form a/o + noun may occur either before or after the possessed noun. This is a general feature of possessive phrases in Māori and is not specific to the GRC. When preposed, genitives are composed of the definite article joined with the a/o + noun construction. This can be seen overtly when the possessed noun is singular, as in (8). When the possessed noun is plural, it is marked by a null morpheme and as such appears to lack a determiner, as in (9). The important point here is that the meaning of the relative clause does not change based on the position of the genitive-marked DP.

In the following section we will focus on the identity of this genitive-marked item in the context of the GRC, looking precisely at which grammatical positions can surface with genitive marking.

### 3 What gets genitive-marked?: not always subjects (Puzzle 1)

The GRC has been described as a method of relativizing upon canonical transitive sentences; its head noun (which is ‘linked’ to the direct object position of the relative clause via abstraction over that position) is modified by a genitive-marked subject (Herd et al. 2011, Otsuka 2010). This is represented schematically in (10), which corresponds to (6) above.

<table>
<thead>
<tr>
<th>Genitive subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>the man of John: [Op1 murdered Øsubj₂ Øobj₁]</td>
</tr>
</tbody>
</table>

‘the man that John murdered’

Interestingly, however, not all clauses that can undergo relativization using the GRC are canonical transitive clauses. Specifically, the GRC can be used to relativize on at least some passive constructions. Such constructions have thus far not been described in the literature on the GRC. In these sentences, the genitive-marked noun does not correspond to the grammatical subject of the relative clause, but rather seems to correspond to an oblique DP (what can be thought of as the by-phrase of a passive). To illustrate, in (11) the GRC appears to relativize on the subject of a passive construction, leaving the oblique DP to receive genitive marking. We show this schematically with English lexical items in (12); (a) has the gap for the relativized noun, and (b) adds the gap for the oblique DP. (An example of a simple passive sentence can be found in (14) in the next section.)

---

3 The distinction between a and o has to do with the real-world relationship between possessors and possessees. In general, a-possession encodes a relationship in which the possessor has control over the possesee, while o-possession encodes the lack of such a relationship (though see e.g. Biggs (1969), Hohepa (1967), Bauer et al. (2003) for more detailed discussion). According to Bauer et al. (2003), the possessor in a GRC configuration takes the form a + noun when the direct object is relativized but may take the form o + noun when an oblique noun associated with an intransitive verb is relativized (570, 577).
(11) Ko tōna ngākau kīhai i wareware ki tana mea
PRED 3SG.POSS heart NEG TAM forget OBJ 3SG.POSS thing
[i kite-a ai Øsubj Øhyphrased hei taonga mōna.
TAM see-CIA PART PREP treasure for.3SG

‘His heart did not forget his thing that he had seen that would be a treasure for him.’ (Grey 2001:174)

(12) a. thing [which1 t1 was seen by him]

b. his thing [which1 t1 was seen (by) Ø]

In the next section, we will consider an analysis of Māori as a split-ergative language, put forth by Pucilowski (2006), and we will show that the asymmetry we have just described goes away under that analysis.

4  Split-ergativity and the GRC

4.1 Māori as a split-ergative language

Māori has traditionally been described as a nominative-accusative language with an active and a passive construction, shown in (13) and (14), respectively. In this section we will be questioning whether these are really cases of active and passive, so to avoid confusion, we will henceforth use the neutral labels “Pattern I” for what has been called the active and “Pattern II” for what has been called the passive, following Clark (1976).

(13) E kai ana ngā tamariki i ngā āporo. Pattern I
TAM eat TAM the.PL children OBJ the.PL apple

‘The children are eating the apples.’ (Bauer et al. 1997:40)

(14) I patu-a te kuri e te tamaiti. Pattern II
TAM hit-CIA the dog AGT the child

‘The dog was hit by the child.’ (Bauer et al. 1997:42)

The status of the Māori passive has long been a topic of interest, as it does not behave how we might expect a passive to behave. For example, the passive in Māori is much more frequent than the passive in other languages, and it occurs particularly often in past tense narratives (for further discussion see Clark 1976). The debate over the passive has led some to propose that Māori is in fact an ergative language (e.g. Sinclair 1976, Modini 1985). In an ergative language, the subject of an intransitive verb and the object of a transitive verb pattern together, while the subject (or agent) of a transitive verb behaves differently. If we were to adopt this analysis for Māori, what we have been calling the passive would actually be the basic sentence type in the language: an active sentence with ergative marking.

Pucilowski (2006) unites the two sides of the debate by proposing that Māori is in fact a split-ergative language. This split, she argues, is based on the transitivity of a clause (see Otsuka 2011 for a similar proposal). Transitivity here is understood in the sense of Hopper and Thompson (1980). It involves several factors and is a property of an entire clause. For Pucilowski, the relevant factors for Māori include the number of participants, the clause’s aspect, the affectedness of the direct object, and the notion of dynamism (as opposed to stativity). At least one of these features must be present in an ergative clause.

Under Pucilowski’s analysis, constructions traditionally called ‘active’ (our Pattern I) have nominative-accusative alignment and are used in sentences with low transitivity, and constructions traditionally called
‘passive’ (our Pattern II) have ergative-absolutive alignment and are used in high transitivity sentences. We can see this difference play out in the domain of aspect by comparing (15) and (16).

(15) E ere ana a Huia i ngā kurī. Pattern I (low transitivity)
TAM tie TAM PERS Huia OBJ the.PL dog
‘Huia was tying up the dogs.’ (Bauer et al. 1997:477)

(16) I here-a e Huia ngā kurī. Pattern II (high transitivity)
TAM tie-CIA AGT Huia the.PL dog
‘Huia tied up the dogs.’ (Pucilowski 2006:44)

The case-marking patterns for the two types of sentences are given in Table 1 (from Pucilowski 2006). The letters A, S, and O refer to the Agent of a transitive clause, the Subject of an intransitive, and the Object of a transitive, respectively, as is standard in descriptions of ergative alignment.

Table 1: Split-ergative marking in Māori

<table>
<thead>
<tr>
<th></th>
<th>Ergative (Pattern II)</th>
<th>Accusative (Pattern I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case</td>
<td>Particle</td>
<td>Case</td>
</tr>
<tr>
<td>A</td>
<td>ergative</td>
<td>e</td>
</tr>
<tr>
<td>S</td>
<td>absolutive</td>
<td>∅</td>
</tr>
<tr>
<td>O</td>
<td>absolutive</td>
<td>∅</td>
</tr>
</tbody>
</table>

Pucilowski’s split-ergative proposal draws on evidence from topicalization, question formation, and relative clauses. All three clause types have relative clause-like structure and allow the use of the GRC in some instances. It should be noted that topicalization and question formation have been analyzed as pseudo-clefts or as having relative clause-like structure in Māori and related languages, so the availability of the GRC in these cases is not surprising: see Potsdam and Polinsky (2011).

According to Pucilowski, topicalization and question formation behave slightly differently from relativization when it comes to the strategies they use for displacement, yet all three show evidence of split-ergativity. We will briefly describe her generalizations for topicalization and question formation below; for examples and further discussion on these constructions, see Pucilowski (2006). We will then go through the data for relative clauses in more detail, drawing attention to a slight modification needed for Pucilowski’s conclusions to capture a fuller range of data.

In topicalization and question formation, Pattern II (high transitivity) clauses use the gap strategy for relativization on S and O, but this strategy is unavailable for A, which uses a construction called the actor-emphatic (see (17) below). Pattern I (low transitivity) clauses, on the other hand, use the gap strategy on S and A, and the GRC on O, where the gap strategy is unavailable.

Table 2: Availability of gap strategy in topicalization and question formation

<table>
<thead>
<tr>
<th></th>
<th>High transitivity (Pattern II)</th>
<th>Low transitivity (Pattern I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>actor-emphatic</td>
<td>✓</td>
</tr>
<tr>
<td>S</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>O</td>
<td>✓</td>
<td>GRC</td>
</tr>
</tbody>
</table>
Relative clauses present a slightly different picture for Pucilowski. She argues that high transitivity clauses have ergative alignment, while low transitivity clauses have a neutral pivot, with all of S, A, and O using the gap strategy of relativization.

**Table 3: Availability of gap strategy in relative clauses**

<table>
<thead>
<tr>
<th>High transitivity (Pattern II)</th>
<th>Low transitivity (Pattern I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A actor-emphatic</td>
<td>✓</td>
</tr>
<tr>
<td>S</td>
<td>✓</td>
</tr>
<tr>
<td>O</td>
<td>✓</td>
</tr>
</tbody>
</table>

The bottom-right cell of Table 3 stands out because it is different from the corresponding cell in Table 2, and different from how we characterized the distribution of the GRC in Section 2. We will therefore walk through some of the data that give us the generalizations in Table 3.

Let us begin by considering the strategies used to relativize the agent of a transitive verb (the first row in Table 3). Because relativization on high transitivity agents is not possible using this method, speakers employ the so-called ‘actor-emphatic’ construction.

The actor-emphatic, found in several Eastern Polynesian languages and described by Bauer et al. (2003) for Māori, has the agent appearing at the beginning of the sentence, accompanied by a possessive preposition. Meanwhile, the patient NP surfaces without the direct object marker and instead has the null marking reserved for subjects. An example is given in (17).

(17) Nā Rewi i whāngai ngā manu.
POSS Rewi TAM feed the.PL bird
‘Rewi fed the birds.’ (Bauer et al. 2003:91)

An example of relativization on the agent of an actor-emphatic construction can be seen in the sentence in (18).

(18) Kua tae mai te kōtiro [nāna i hoki mai ngā whurutu].
TAM arrive hither the girl POSS.3SG TAM buy hither the.PL fruit
‘The girl who bought the fruit has arrived.’ (Bauer 1982:324)

Given Pucilowski’s analysis, we expect that relativization using the actor-emphatic should be in complementary distribution with the gap strategy. Crucially, this complementary distribution should be based on the transitivity of the clause. This prediction is in fact supported by the judgments of Pucilowski’s informants. She notes that, whereas relativization on a highly transitive underlying sentence uses the actor-emphatic construction, as in (18), the simple gap strategy is preferred for sentences lower on the transitivity scale (19).

(19) Ka tū anō taua koroheke [i arahi mai rā Øsubj].
TAM stand again DEM old.man TAM lead hither PROX
OBJ a Puhihuia.
‘The old man who had led Puhihuia here stood up again.’ (Bauer et al. 1997:566)
Now, moving to the S row of the Table 3, we expect to see the gap strategy available for the subjects of all intransitive verbs, and this is exactly what we find.

(20)  
\[
\begin{align*}
\text{te} & \quad \text{tamaiti} \quad [i \quad \text{mate} \quad \emptyset_{\text{subj}}]. \\
\text{the \ child} & \quad \text{TAM} \quad \text{die} \\
\text{‘the child that died’ (Orbell 1968:8)}
\end{align*}
\]

Finally, we can examine the behavior of direct objects (the final row in Table 3). Direct objects in Pattern II (high transitivity) clauses can be relativized using the gap strategy, as in (21).

(21)  
\[
\begin{align*}
I & \quad \text{waiata a} \quad \text{Inia} \quad i \quad \text{te} \quad \text{waiata} \quad [i \quad \text{titoa} \quad e] \\
\text{TAM} & \quad \text{sing} \quad \text{PERS} \quad \text{Inia} \quad \text{OBJ} \quad \text{the \ song} \quad \text{TAM} \quad \text{compose.CIA \ AGT} \\
\text{Alfred} & \quad \text{Hill}.
\end{align*}
\]

Alfred  Hill

‘Inia sang the song that Alfred Hill composed.’ (Bauer 1982:312)

The data for the bottom-right cell of the table is slightly more complicated. Pucilowski is correct in pointing out that some Pattern I sentences do in fact allow extraction using the gap strategy, specifically those sentences with imperfective aspect, such as (22). It is from this data that she concludes that low transitivity sentences have a neutral pivot when it comes to relative clauses. However, it is also true that the gap strategy is unavailable for the objects of many Pattern I sentences, such as (23).

(22)  
\[
\begin{align*}
I & \quad \text{hoko mai ana a} \quad \text{ia i} \quad \text{ngā kūmara} \quad [e \quad \text{whakatipu}] \\
\text{TAM} & \quad \text{buy} \quad \text{here} \quad \text{TAM} \quad \text{PERS} \quad \text{3SG} \quad \text{OBJ} \quad \text{the.PL kumara} \quad \text{TAM} \quad \text{make.grow} \\
\text{ana a} & \quad \text{Hata} \quad \emptyset_{\text{obj}}]. \\
\text{TAM} & \quad \text{PERS} \quad \text{Hata} \\
\text{‘He buys the kumara Hata grows.’ (Bauer 1982:316)}
\end{align*}
\]

(23)  
\[
\begin{align*}
*I & \quad \text{hoko mai ia i} \quad \text{te} \quad \text{whare} \quad [i \quad \text{hanga a} \quad \text{Hata} \quad \emptyset_{\text{obj}}]. \\
\text{TAM} & \quad \text{buy} \quad \text{hither he} \quad \text{OBJ} \quad \text{the \ house} \quad \text{TAM} \quad \text{build} \quad \text{PERS} \quad \text{Hata} \\
\text{Intended: ‘He bought the house Hata built’ (Bauer 1982:310)}
\end{align*}
\]

Both (22) and (23) are Pattern I sentences, yet they differ in whether the gap strategy is possible. Assuming (following Hopper and Thompson 1980) that perfective clauses are higher on the transitivity scale than imperfective clauses, the data above suggest that there are more gradations of transitivity than a simple “high/low” dichotomy might suggest. Still, it seems clear that transitivity plays a significant role in determining the availability of relativization strategies in Māori.

4.2 The GRC under a split-ergative analysis

Under the split-ergative analysis, the identity of the genitive-marked noun receives a more natural characterization. Pattern II clauses use the gap strategy on S and O, but not A, while Pattern I clauses use the gap strategy on S and A, and not on O (with the caveat discussed in Section 4.1). These generalizations are summarized in Table 4.
Table 4: Availability of gap strategy.

<table>
<thead>
<tr>
<th>High transitivity (Pattern II)</th>
<th>Low transitivity (Pattern I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>S</td>
<td>✓</td>
</tr>
<tr>
<td>O</td>
<td>✓</td>
</tr>
</tbody>
</table>

Now, if we analyze Pattern II sentences (previously ‘passive’) as ergative-absolutive constructions, then (11), repeated from above, is less surprising. Specifically, the puzzle regarding the grammatical position of the genitive-marked noun goes away. Under the split-ergative analysis, the genitive-marked noun in (11) is associated with the subject of the relative clause, while the object position is relativized, just like in a Pattern I relative clause that uses the GRC. This is schematized in (24a), which contrasts with the passive analysis in (24b), where the relativized position is the subject.

(11) Ko tōna ngākau kīhai i wareware ki tana mea
PRED 3SG.POSS heart NEG TAM forget OBJ 3SG.POSS thing
[i kite-a ai Øsubj Øhyphrase] hei taonga mōna.
TAM see-CIA PART PREP treasure for.3SG

‘His heart did not forget his thing that he had seen that would be a treasure for him.’ (Grey 2001:174)

(24) a. Ergative analysis: his thing [which t1 saw t1]

b. Passive analysis: his thing [which t1 was seen (by) Ø]

Adopting the split-ergative analysis allows us to more naturally characterize the GRC’s distribution as that of a relativization strategy for structural positions that are less accessible (in the sense of Keenan and Comrie 1977) than subjects. Characterizing the GRC in this way removes the need to explain the availability of the GRC for a particular class of subjects.

5 What position gets relativized? (Puzzle 2)
We now turn to the question of which grammatical positions can undergo relativization using the GRC. As we saw above, there is an extraction restriction on the direct object of Pattern I sentences: these sentences cannot be relativized with the gap strategy (unless they have imperfective aspect). In light of this restriction, the GRC has been described as a kind of “rescue strategy” allowing one to circumvent the extraction restriction and relativize on the direct object of a canonical transitive verb, as in Herd et al. (2011), which assumes that the underlying object in the GRC moves out of the object position before being extracted.

While it is true that the GRC allows relativization on direct objects, framing it as a rescue strategy obscures the fact that the GRC has a wider distribution than just relativization on direct objects. In this section we will sketch this wider distribution.

5.1 The GRC as a rescue strategy
Herd et al. (2011) proposes that, next to passivization (25) and the Māori actor-emphatic construction (26), both of which can be analyzed as having an underlying direct object that surfaces in subject position (Herd et al. 2011), the GRC in Polynesian is yet another way to get around the well-known extraction restriction on relative clause formation in Polynesian languages, i.e., the impossibility of relativizing directly upon the direct object position. In the passive and actor-emphatic sentences below, the relevant arguments are in bold. These arguments appear with the null marking usually associated with subjects.
Herd et al. (2011) takes the actor-emphatic construction to involve an underlying object moving to the subject position, similar to a passive. Then, assuming that both the passive and the actor-emphatic allow one to relativize on an underlying direct object by placing the object in a derived subject position, Herd et al. (2011) argues that the GRC also involves a derived subject position (27). The analogous, and more familiar, passive construction is given in (28). The authors put forth an accessibility-based analysis in which direct objects cannot undergo relativization, so the relativization site in the GRC must therefore be some other position.

In the next section we will show that characterizing the GRC as a rescue strategy is somewhat misleading, as that would imply a narrower distribution of the construction than we actually see.

5.2 Wider distribution of the GRC: other relativized positions
Under the accessibility-based analysis proposed in Herd et al. (2011), the GRC is characterized as a rescue strategy that is used in Pattern I sentences when relativization on the direct object position is unavailable. We will now show that the GRC has a wider distribution than is straightforwardly predicted by that analysis. In particular, we will show that along with relativizing the direct objects of Pattern I sentences, the GRC can be used to relativize Pattern II objects (as discussed in Section 4.2), as well as oblique DPs and the objects of experience verbs (Bauer et al. 2003:577).

First, as we have already seen, the GRC is compatible with relativization on the object position in a Pattern II (previously ‘passive’) sentence (11, repeated from above). It should be noted that in a nominative-accusative analysis of Māori, where such sentences are indeed passives, the relativized position in (11) would be the subject. This is unexpected if the GRC is characterized purely as a rescue strategy. In other words, there is no extraction restriction here, so we lose the hypothesized motivation for the GRC.

We now present two more cases where the GRC cannot be characterized as a rescue strategy, as it is used to relativize on positions other than the objects of canonical transitive sentences. Crucially, these positions allow relativization using other (non-GRC) strategies, and so they do not need to be “rescued.”
The first such case is that of oblique DPs, which permit relativization using the GRC (Bauer et al. 2003:577). In Māori, oblique DPs are DPs that are introduced by a preposition (usually ki). Oblique DPs can undergo direct relativization with a strategy other than the GRC (in the case of (29), with a resumptive pronoun).

(29) Ko Tamahae te tamaiti fi mau nei i a ia te tarakihi.
3SG the child TAM be.caught PROX CAUS PERS the tarakihi

‘Tamahae is the child by whom the terakihi was caught.’ (Bauer et al. 2003:54)

The example in (30) shows an oblique DP relativized using the GRC, schematized in (31).

(30) I hoe mai hoki te waka rā i muri i te kōtiro
TAM paddle hither also the canoe PROX at behind at the girl
rā i te wā ōna fi rere rā ə subj ki te
PROX at the time 3SG.POSS TAM jump PROX to the
wai ə oblq.

‘The canoe had also rowed up behind the girl at the same time when she had jumped into the water.’ (Bauer et al. 1997:569)

(31) at her time [when, ə subj 2 had jumped into the water ə oblq 1]
‘at the time when she had jumped into the water’

Given that there is no restriction against relativizing oblique DPs, and the GRC is not the only strategy that can target the oblique position, a derived subject analysis is not motivated with respect to relativization on obliques.

The rescue strategy characterization of the GRC encounters a similar problem with relativization on the objects of experience verbs. Experience verbs in Māori are verbs that take two arguments, with neither passing all of the diagnostics for direct objects in the language (Bauer et al. 2003). These verbs generally have their second argument marked with ki. Examples include verbs such as pirangi ‘want’, mōhio, ‘know’, and wareware ‘forget’ (32).

(32) Kua wareware au ki tana ingoa.
TAM forget 1SG OBJ 3SG.POSS name

‘I have forgotten his name.’ (Harlow 2007:109)

4 An exception is the verb kite (“see”), which is marked with i (usually a direct object marker) even though it patterns syntactically with experience verbs (see Bauer et al. (2003) for more on experience verbs). Conversely, some direct objects of canonical transitive verbs are marked with ki, although i is much more common. This is not important for our purposes, except to say that the identity of the preposition does not determine whether a verb is a canonical transitive verb or an experience verb.
Further evidence for this distinction is that the ki-marked arguments of experience verbs do not behave like those of canonical transitives when it comes to relativization. Unlike those of canonical transitives, objects of experience verbs can use the gap method, just like subjects, as we can see in (34).

(34) *I tūtaki a ia ki te tamaiti [i mōhio a Rewi ∅obj]*.
    TAM meet PERS 3SG OBJ the child TAM know PERS Rewi
    ‘He met the child that Rewi knew.’ (Bauer et al. 2003:57)

Unlike subjects, however, objects of experience verbs can also be relativized using the GRC (35), schematized in (36).

(35) *Ko tēnei te whare a Hata [i pīrangi at ∅subj ∅obj]*.
    PRED this the house of Hata TAM want PART
    ‘This is the house that Hata wanted.’ (Bauer et al. 1997:569)

(36) the house of Hata [Op₁, ∅subj, wanted ∅obj]

Again, there is no motivation for the GRC to involve a derived subject position, as it is acceptable to relativize on the object of an experience verb with the gap method.

In sum, the GRC has a wider distribution than just direct object relativization, which is surprising under a pure rescue strategy analysis, and which suggests that the distribution of the GRC is better characterized as a strategy that is available whenever a noun other than the agent/experiencer is being relativized (because the agent/experiencer is reserved for the genitive-marked position). Further research should help determine whether such a characterization can shed light on the syntactic mechanisms underlying the GRC. A few such questions are put forth in the following section.

6 Summary and next steps

Our goal was to characterize the distribution of the GRC by focusing on (i) what gets genitive-marked, and (ii) which position gets relativized. We showed that under a nominative-accusative analysis of Māori, the genitive-marked noun is not limited to subjects but is also available on the oblique DP in a passive. We then demonstrated that this apparent irregularity disappears under the split-ergative analysis of Māori, which Pucilowski (2006) argues for with independent evidence.

As for the question of which position gets relativized, we showed that relativization is not restricted to direct objects, and that the GRC is actually compatible with a wider range of relativized positions, some of which do not lend themselves to a characterization of the GRC simply as a rescue strategy.

As we said in Section 1, we need to leave for future research what the mechanism is for generating genitive subjects outside of their associated relative clauses. To understand the mechanism involved, we need to keep in mind questions such as (i) how we can capture the fact that genitive-marked arguments are always interpreted as subjects/agents in the relative clauses; (ii) whether there is any movement relation involved, and if so, what type of movement; (iii) what, if any, the connections are between the GRC and the actor-emphatic, and the Samoan Possessor-Agent-Goal (PAGO) constructions (Homer 2009).
References

A PRESENTATIONAL CONSTRUCTION IN INDONESIAN

Carly J. Sommerlot
University of Texas at Arlington
carly.sommerlot@mavs.uta.edu

Abstract
Despite being a canonically SVO language, Indonesian permits verbs in sentence-initial position in some contexts. This study investigates a previously undescribed type of verb-initial construction that co-occurs with the application of two morphemes: the intransitive prefix ber- and the suffix -nya. I argue that this ber-V-nya construction, unlike other verb-initial constructions in Indonesian, is analogous to the presentational-there construction in English as described by Aissen (1975), as they share several characteristics. I further argue that the same syntactic structure offered for other verb-initial constructions in Indonesian (Chung 2008) can be applied to these ber-V-nya constructions.

Keywords: syntax, Indonesian, presentational, clause structure, verb-initial

ISO 639-3 codes: ind

1 Introduction
While the neutral word order of clauses in Indonesian is well-attested to be SVO, there has been much research highlighting non-canonical word orders and their derivation (Chung 1978; Kaswanti Purwo 1989; Chung 2008; Sneddon et al 2010). It seems that, under certain circumstances, verbs can occur sentence-initially. These ‘certain circumstances’ have been argued to be related to information packaging and phonology, with the purpose of highlighting the predicate instead of the subject (Chung 1978; Kaswanti Purwo 1989). However, there still remains many questions about the precise triggers for this non-canonical word order.

This study seeks to investigate one particular construction that demonstrates verb-initial word order, a construction I will refer to as ber-V-nya. In particular, I will focus on features that are unique to this construction and how these features parallel those of the presentational-there construction in English (Aissen 1975). I will further discuss some morphological restrictions on these ber-V-nya constructions, noting that they must co-occur with two morphemes, ber- and -nya. These two morphemes have received some attention over the past years, but still seem to escape full description. The verbal prefix ber- has generally been treated as a marker of intransitivity (Butar-Butar 1976; Dardjowidjojo 1978; Sie 1989; Vamarasi 1999), but more recent work has attempted to either unpack this further (Udayana & Beavers 2013), or disprove this assumption altogether (Fortin & Soh 2014). The suffix -nya, which can attach to both nouns and verbs, has been argued to have a multitude of functions, including marking the 3rd person genitive, definiteness, nominalization, and evidentiality (Englebretson 2003; Yap 2011; Arka 2011; Grangé 2015). What has not been previously discussed is the interaction of these two morphemes - what functions do these two morphemes have when paired together?

Section 2 introduces the ber-V-nya construction and outlines its features, while also discussing how it fits into the current literature. Section 3 discusses parallels between ber-V-nya constructions and presentational-there sentences in English, arguing that these ber-V-nya constructions are actually a type of presentational. Section 4 discusses potential analyses, utilizing three diagnostics that determine the position of the subject in these ber-V-nya constructions to argue for using Chung’s (2008) analysis for verb-initial constructions in Indonesian. Section 5 concludes with lingering questions and future directions.

1 I would like to first and foremost thank the Serworwora family for serving as language consultants for the past few years on this project. Any mistakes in interpretation are my own. I’d also like to thank Joey Sabbagh, Laurel Stvan, and Suwon Yoon for their valuable feedback on previous versions of this project. Lastly, I’d like to thank the audience at AFLA 25 and two anonymous reviewers for their insightful comments that contributed to improving this paper.
Features of *ber-V-nya*

The data to be provided in this paper is novel data of the type *ber-V-nya*. In the following sections, I will present data of this type, providing description only, and highlighting notable features of these constructions. An analysis of this data will be presented in Section 4.

2.1 Non-canonical word order

The most notable feature of these *ber-V-nya* constructions is that they do not follow canonical SVO word order. Verbs, including those prefixed with *ber-*, normally occur after the subject, as in (1a), and cannot occur preverbally as in (1b):

(1a) *Mereka ber-kelahi meng-guna-kan sepotong kayu.*
3PL BER-fight AV-use-CAUS piece wood

‘They fight with/using pieces of wood.’

(1b) *Ber-kelahi mereka meng-guna-kan sepotong kayu.*
BER-fight 3PL AV-use-CAUS piece wood

‘They fight with/using pieces of wood.’

With the addition of -nya, however, the affixed verb must be fronted to the sentence-initial position as given in (2b); leaving the affixed verb in its canonical position after the subject is ungrammatical, as in (2a).

(2a) *Mereka ber-kelahi-nya meng-guna-kan sepotong kayu.*
3PL BER-fight-NYA AV-use-CAUS piece wood

‘They fight with/using pieces of wood.’

(2b) *Ber-kelahi-nya mereka meng-guna-kan sepotong kayu.*
BER-fight-NYA 3PL AV-use-CAUS piece wood

‘They fight with/using pieces of wood.’

2.2 Necessity of additional information

Another notable feature of these constructions is the necessity of additional information that describes the action occurring. Compare the grammatical sentence in (2a) above to the sample in (3), which lacks the adjunct phrase with the additional information about what was being used during the fight, and instead only contains the verb in the sentence-initial position followed by the subject.

(3) *Ber-kelahi-nya mereka.*
BER-fight-NYA 3PL

‘They fight.’

In (3), the sentence becomes ungrammatical when *menggunakan sepotong kayu* ‘with/using pieces of wood’ is eliminated. Importantly, the sentence given in (3) would be grammatical if we reverted back into a more canonical construction with SV order, and removed -nya:

(4) *Mereka ber-kelahi.*
3PL BER-fight

‘They fight.’
Only when the verb is affixed with -nya and fronted does this adjunct providing additional information seem necessary. This adjunct phrase can be locative (5a-b), temporal (6a-b), or a reason (7a).

(5a) *Ber-enang-nya ikan itu di laut.
BER-swim-NYA fish that in sea
‘The fish swims in the sea.’

(5b) Ber-nyanyi-nya mereka di gereja.
BER-sing-NYA 3PL in church
‘They sing in church.’

(6a) Ber-main-nya dia dengan komputer sampai tengah malam.
BER-play-NYA 3SG with computer until middle night
‘He played with the computer until midnight.’

(6b) Ber-sin-nya saya waktu sedang mem-baca.
BER-sneeze-NYA 1SG while PROG AV-read
‘I sneezed while reading.’

(7) Ber-sin-nya saya karena bau bensin.
BER-sneeze-NYA 1SG because smell gasoline
‘I sneezed because I smelled gasoline.’

Each of the sentences given in (5)-(7) becomes ungrammatical when the adjunct is removed, like the sentence given in (3). Such a restriction seems to indicate that these ber-V-nya constructions necessitate additional information that describes the action occurring; in particular, they necessitate information that is not required without -nya, as indicated by the acceptability of the sentence given in (4). However, this additional information cannot occur in the form of a DP argument. Contrast (5b) above with (8) below, where the verb bernyanyinya ‘sing’ has now taken the object lagu-lagu ‘songs’ and is consequently ungrammatical.

(8) *Ber-nyanyi-nya mereka lagu-lagu di gereja.
BER-sing-NYA 3PL song-RED in church
‘They sing songs in church.’

Verbs prefixed with ber- can occasionally take direct objects when not participating in ber-V-nya constructions, which will be discussed in Section 2.4
3 Co-occurrence with aspect markers

Some verbs that participate in these ber-V-nya constructions can co-occur with the aspect markers sedang ‘progressive’ and sudah ‘perfect’. These aspect markers can only occur preverbally, like in (9a), placing them in the sentence-initial position. Occurring after the verb is ungrammatical, as exemplified by (9b).

(9a) Sedang ber-nyanyi-nya mereka di gereja.
PROG BER-sing-NYA 3PL in church
‘They are singing in church.’

(9b) *Ber-nyanyi-nya sedang mereka di gereja.
BER-sing-NYA PROG 3PL in church
‘They are singing in church.’

The progressive aspect marker sedang is used above in (9), but this same pattern applies for the perfect aspect marker sudah as well, as seen in (10a-b).

(10a) Sudah ber-main-nya dia dengan computer.
PERF BER-play-NYA 3SG with computer
‘He played with the computer.’

(10b) *Ber-main-nya sudah dia dengan computer.
BER-play-NYA PERF 3SG with computer
‘He played with the computer.’

2.4 Can we rely on previous analyses to account for this data?

Given the features outlined above, how can we account for these constructions? One possibility is to consider previous analyses of the two morphemes involved, ber- and -nya.

2.4.1 Previous analyses of ber-

As noted earlier, ber- has generally been assumed to be an intransitivizer. However, some verbs prefixed with ber- can take what on the surface appears to be an object, as in (11a), where bernyanyi ‘sing’ takes the object lagu-lagu ‘songs’, and in (11b), where bermain ‘play’ takes the object permainan komputer ‘computer games’:

---

2 It should be noted that these are generally categorized differently than auxiliaries. In Indonesian, modals are usually categorized as auxiliaries. At least some of these are compatible with ber-V-nya as well:

(1) Perlu berenang-nya ikan untuk tetap hidup.
must BER-swim-NYA fish for still life
‘Fish must swim to stay alive.’

It is uncertain at this point whether all modals are compatible with ber-V-nya. This compatibility may also depend on a general compatibility with ber- verbs and is thus beyond the scope of this paper.

3 Note that some ber- verbs seem to resist aspect marking in general, but this occurs in canonical constructions as well.

4 Sneddon et al (2010) refers to these as both ‘temporal markers’ and ‘aspect markers’, as they indicate completeness or continuation of an action. They are labeled as separate from adjuncts of time, such as tadi pagi ‘this morning’. However, the role of these markers is not completely clear. While Sneddon defines sudah as indicating that an action has occurred or a state has been achieved, Grangé (2013) and Kaswanti Purwo (1984) note that sudah also indicates the subject’s subjectivity and therefore also expresses modality. It would be interesting to explore further the compatibility of ber-V-nya with sudah and to see if this relates to previous analyses of -nya as a marker of evidentiality; this is an issue I hope to pursue in future research.

5 Historically, ber- is a cognate with the antipassive/unergative prefix mag- in Tagalog. In Tagalog, objects in antipassives are typically required to be non-specific; this could also be the case in Indonesian in ber- constructions but requires further investigation. I would like to thank an anonymous reviewer for pointing this out.
Papers from AFLA 25 – Sommerlot

(11a) Mereka ber-nyanyi lagu-lagu di gereja.
3PL BER-sing song-RED in church
‘They sing songs in church.’

(11b) Dia ber-main permainan komputer sampai tengah malam.
3SG BER-play games computer until middle night
‘He played computer games until midnight.’

Such constructions might call into question ber-’s function as a intransitivizer. Of particular interest here, however, is that constructions like (11a-b) cannot participate in the non-canonical word order of ber-V-nya, as exemplified in (12a-b) and (13a-b) below.

(12a) *Ber-nyanyi-nya mereka lagu-lagu di gereja.
BER-sing-NYA 3PL song-RED in church
‘They sing songs in church.’

(12b) *Ber-nyanyi-nya lagu-lagu mereka di gereja.
BER-sing-NYA song-RED 3PL in church
‘They sing songs in church.’

(13a) *Ber-main-nya dia permainan komputer sampai tengah malam.
BER-play-NYA 3SG games computer until middle night
‘He played computer games until midnight.’

(13b) *Ber-main-nya permainan komputer dia sampai tengah malam.
BER-play-NYA games computer 3SG until middle night
‘He played computer games until midnight.’

The ungrammaticality of (12a-b) suggests that ber-V-nya is limited to intransitives. It has been argued that ber- verbs occurring with objects as in (11) are actually instances of incorporation; the post-verbal nominals are not true direct objects, but instead have been incorporated into the verb (Vamarasi 1999; Dardjowidjojo 1978; Butar-Butar 1976; Sie 1989). Such an analysis might actually account for the ungrammaticality of (12) and (13), as -nya would interrupt the prosodic phrase between the verb and its post-verbal nominal, but this remains a question for future research.

It could be hypothesized that ber- is merely a coincidence in ber-V-nya constructions. In the examples given in the sections above, -nya seems to be the main trigger of most of the features, including the non-canonical order. I argue that this is not the case, as ber- must occur in these ber-V-nya constructions. Evidence comes from so-called ‘optional’ ber- verbs (Fortin and Soh 2014), which can occur both with or without ber-affixed7. Crucially, while these verbs can occur with or without ber- in canonical SVO sentences, only when affixed with ber- can they occur in ber-V-nya constructions. In (13), the verb occurs with ber-, and both the SVO sentence and the ber-V-nya construction are grammatical. In (14), however, the verb is bare, and only the SVO sentence is grammatical (in (14a)) while the ber-V-nya construction is not possible (in (14b)).

---

6 Thanks for an anonymous reviewer for suggesting the inclusion of this data.
7 Most native speakers claim that the only difference between these two is a formality difference; while both bermain and main are possible, bermain is more formal than main.
This suggests that *ber*- is not only functioning as an intransitivizer in *ber-V-nya* constructions, and that *ber*- actually has more complex functions that previously thought. It is not possible, then, to rely on previous analyses of *ber*- to account for *ber-V-nya* constructions.

2.4.2 Previous analyses of *-nya*

Another potential direction is to consider previous analyses of *-nya*. These *ber-V-nya* constructions do seem reminiscent of Grangé’s (2015) subordinated NPs; both of these constructions a) occur sentence-initially, and b) consist of a verb affixed with *-nya* followed by an NP.

Notably, both Grangé’s subordinated NPs and the *ber-V-nya* constructions outlined here occur sentence-initially. One crucial difference, however, is that Grangé’s *verb-nya* sentences can optionally front, while the *ber-V-nya* sentences must be fronted to be grammatical. Grangé notes that some stative verbs may occur as the head of an object NP, as in (15) below.

Governor Jakarta AV-regret-APPL less-NYA supply-NOMI gas
‘The Governor of Jakarta regrets the shortage of natural gas supply.’ (Grangé 2015:23)

The *verb-nya* construction in (15) is *kurangnya pasokan gas* ‘shortage of natural gas supply’. This does not occur sentence-initially, as it is the object of the sentence. This is not possible in *ber-V-nya* constructions, as they are limited to intransitives.

The second similarity between Grangé’s data and the data presented above is the pattern of having a verb affixed with *-nya* followed by an NP. The complement NP in Grangé’s data originates from a demoted subject. Grangé provides the following in (16) as evidence of this:

(16) Terumbu karang mati. → matinya terumbu karang
coral reef die die-NYA coral reef
‘Coral reefs die.’
‘the death of coral reefs’
Literally: ‘the (fact that) die coral reefs’
(Grangé 2015:55)

In (16), the subject *terumbu karang* ‘coral reef’ becomes the complement NP of the subordinated NP *matinya* ‘death’. This is characteristic of every example given and is described by Grangé as a necessary component of these constructions. In Section 4 I provide evidence that the subject in *ber-V-nya* constructions is actually located in the spec,TP, which differs from the analysis given by Grangé. Furthermore, while these
subordinated NPs require a complement NP, *ber-V-nya* constructions require an adjunct with additional information describing the action to be grammatical. No such requirement exists in Grangé’s analysis.

Further differences indicate that these *ber-V-nya* constructions do not fit into any previous analyses. Grangé’s subordinated NPs readily take verbal negation. *ber-V-nya* constructions cannot be negated, using the verbal negator *tidak* in (17b) or the nominal negator *bukan* in (17c).

(17a) *Ber-temu-nya kami di Jakarta.*
BER-meet-NY 1PL in Jakarta
‘We meet in Jakarta.’

(17b) *Tidak ber-temu-nya kami di Jakarta.*
NEG BER-meet-NY 1PL in Jakarta
‘We did not meet in Jakarta.’

(17c) *Bukan ber-temu-nya kami di Jakarta.*
NEG BER-meet-NY 1PL in Jakarta
‘We did not meet in Jakarta.’

(17a) shows a grammatical *ber-V-nya* construction, with no negation. When negated, using either negator available in Indonesian, it becomes ungrammatical, as indicated by (17b) with *tidak*, and (17c) with *bukan*. Verbal negation is possible in Grangé’s analysis, as exemplified in (18) below.

(18) *Tidak masuk-nya anggota para KPU itu karena khawatir.*
NEG enter-NYA member DET PRN that because worry
‘(The fact that) the Election Board members did not attend is because they were worried.’

(Grangé 2015:51)

It seems that these *ber-V-nya* constructions cannot, then, be categorized as ‘subordinated NPs’ under Grangé’s analysis. These constructions also differ from Arka’s (2011) and Englebretson’s (2003) nominalizations, as they have significant verbal properties (such as being able to take aspect markers, and not being able to co-occur with the nominal negator, as evidenced by (17c)). Crucially, Arka notes that nominalizations using -*nya* cannot co-occur with any aspect markers which is not a feature of the data presented here.

One interesting note, however, is that many of these analyses of -*nya* have attributed it to a marker of evidentiality. A unified analysis of this affix, then, might play upon those previous findings; while I believe a new analysis needs to be proposed for these *ber-V-nya* constructions, I also believe that that new analysis should fall under the umbrella of evidentiality to fit with these previous analyses. The analysis I propose below, for *ber-V-nya* constructions as a type of presentational construction, could potentially fall within that category.

3 *ber-V-nya* as a type of presentational-*there* construction

Given the facts and features outlined above, I am proposing that these *ber-V-nya* constructions do not fit into any previously described functions of either of these affixes; instead, the interaction of these two morphemes gives rise to a new type of construction: a type of presentational-*there* construction, analogous to those in English, provided in (19).

(19a) There stands in the corner of the room an old file cabinet.

(19b) Suddenly there ran out of the bushes a grizzly bear.

(Aissen 1975:1,4)
In the following subsections, I will present evidence that ber-V-nya constructions are equivalent to presentational-there constructions in Indonesian, using diagnostics detailed by Aissen (1975). This evidence comes from similarities between the two constructions, which include root restriction, resistance to negation, the necessity of additional information, and markedness, each of which will be detailed below.

3.1 Root restriction
Aissen argues that presentational-there constructions in English are root transformations only. Her evidence for this comes from their inability to be the complement of a non-assertive predicates, as in (20), and indirect questions, as in (21).

(20a) He says that there stepped out in front of his car a pedestrian.
(20b) *The driver regrets that there stepped out in front of his car a pedestrian.
(21) *We were wondering whether there walked into your class today a pack of dogs. (Aissen 1975:5-6)

This is also the case for ber-V-nya constructions, although these take this a step further and are disallowed in all embedded contexts. For example, these constructions cannot occur as the complement to the verbs lihat ‘see’, dengar ‘hear’, and kira ‘think’. When within one of these embedding contexts, they must instead remain in their canonical position without -nya. (22) shows the root clause without -nya and with -nya, while (23) shows them within the embedded predicate saya lihat ‘I see’, again with -nya and without -nya.

(22a) Anak itu ber-main dengan main-an.
Child that BER-play with play-NOMI
‘The child plays with the toys’.

(22b) Ber-main-nya anak itu dengan main-an.
BER-play-nya child that with play-NOMI
‘The child plays with the toys’.

(23a) Saya lihat anak itu ber-main dengan main-an.
1SG see child that BER-play with play-NOMI
‘I see the child plays with the toys’.

(23b) *Saya lihat ber-main-nya anak itu dengan main-an.
1SG see BER-play-nya child that with play-NOMI
‘I see the child plays with the toys’.

As a root clause, (22b) with -nya affixed to the ber- verb is completely grammatical. However, when this same construction occurs in an embedding context, like in (23), it becomes ungrammatical, as exemplified by (23b). This also occurs with other embedding verbs, such as dengar ‘hear’ in (24), and kira ‘think’ in (25) below.

(24a) Saya dengar kedua sahabat itu ber-temu di Dallas.
1SG hear both friend that BER-meet in Dallas
‘I heard that the two friends met in Dallas’.
Saya dengar ber-temu-nya kedua sahabat itu di Dallas.
‘I heard that the two friends met in Dallas’

Saya kira mereka ber-henti di rest area.
‘I think they stopped at a rest area’

Saya kira ber-henti-nya mereka di rest area.
‘I think they stopped at a rest area.’

Such evidence indicates that, like the English presentational-there, ber-V-nya constructions are limited to root clauses only.

3.2 Resistance to negation
Aissen notes that English presentational-there constructions never allow sentential negation.

*There never stepped out in front of my car a pedestrian.

*I don’t believe there ran out of the woods a grizzly bear. (Aissen 1975:46a-b)

She further notes that constructions of this type can only take contrastive constituent negation on the old subject, as exemplified in (27).

There ran into our campsite not a grizzly bear but a panda bear. (Aissen 1975:48a-b)

Indonesian ber-V-nya constructions pattern similarly. They resist sentential negation, even when it occurs preverbally (like aspect markers do).

Ber-temu-nya tidak kami di Jakarta.
‘We did not meet in Jakarta.’

Tidak ber-temu-nya kami di Jakarta.
‘We did not meet in Jakarta.’

Furthermore, these constructions can take constituent negation using the negator bukan within its phrase. This is given in (29).

Ber-kelahi-nya mereka men-guna-kan bukan sepotong kayu.
‘They fight using not pieces of wood.’
This does differ from the English examples given in (26), as the constituent negation occurs within the additional information necessitated by this construction instead of on the post-verbal subject. Nonetheless, ber-V-nya constructions pattern similarly in terms of negation as English presentational-there sentences.

3.2 Necessity of additional information

One of the key similarities between presentational-there constructions and ber-V-nya constructions is the necessity of additional information. In presentational-there constructions, a locative is necessarily included, and this locative occurs on the surface in a position higher than the old subject. Removing this information is not possible:

(30a) There jumped out of the bushes a man.

(30b) *There jumped a man.

The primary focus of this construction, then, is not the subject but the locative information added. The promotion of both the verb and the locative above the subject serves to place focus on this information. As noted earlier, ber-V-nya constructions also necessitate additional information – which can occur as a locative phrase, but it can also occur as a temporal or other type of phrase. Furthermore, these constructions induce non-canonical word order, seemingly placing focus on the verb, not the subject, of the sentence.

(31a) Ber-diri-nya saya di depan kelas.
      BER-stand-NYA 1SG in front class
      ‘I stand in the front of the class.’

(31b) *Ber-diri-nya saya.
      BER-stand-NYA 1SG
      ‘I stand.’

(31c) *Saya ber-diri-nya di depan kelas.
      1SG BER-stand-NYA in front class
      ‘I stand in the front of the class.’

Only (31a) is grammatical, as it has verb-initial ordering as well as a locative phrase. While ber-V-nya constructions do not have the necessary additional phrase moving prior to the subject, they induce non-canonical word order, and the necessity of that additional information only in these types of constructions suggests that such information is the focus of these phrases.

3.3 Markedness

The last similarity between these two constructions is markedness. While Aissen does not specifically discuss this, it seems to be the case that presentational-there constructions are significantly more marked than the more ‘standard’ existential in English; many speakers find such examples as There jumped out of the bushes a man to be slightly odd, and they certainly do not occur with nearly the frequency as the more standard there existential.

This is also the case for Indonesian ber-V-nya constructions. While speakers seem to have strong intuitions about what is allowed and disallowed in these constructions, they do find them to be marked, stating that these were not sentences they would hear on a daily basis. Furthermore, these constructions are considered to be highly informal and would likely not be written in any form. If we assume these are the Indonesian equivalent to presentational-there constructions, this finding would not be odd, but expected.
3.4 Differences between presentationals and ber-V-nya

It should be noted that, while having significant similarities, there are some differences between English presentationals and ber-V-nya constructions. One, ber-V-nya constructions frequently allow pronouns in the postverbal subject position. In English, this is not possible, as seen in (32) below.

(32a) *There jumped out of the bushes me.
(32b) *There stands in the corner him.

The presentational in English contrasts with Indonesian ber-V-nya as the subject is typically indefinite, as it is argued to be discourse-new information (Ward and Birner 2001). However, while pronouns are not possible, definite subjects are sometimes possible in English presentationals, as exemplified in (33) below.

(33) Suddenly there jumped out of the bushes the grizzly bear we saw earlier.

This is also the case for Indonesian. The postverbal subject can be both a pronoun and modified by the determiner itu. This is discussed more in Section 4.3 below. It could be the case that languages differ in regards to what types of definites they allow in presentationals, with Indonesian allowing pronouns, while English does not.

A second difference is in the position of the additional information. In English, a PP occurs between the verb and the postverbal subject, as seen in (30a), where the PP out of the bushes occurs between the verb jump and the subject a man. In Indonesian, the additional information must occur after the postverbal NP, as seen in (34a). If the additional information occurs between the verb and postverbal NP, the construction becomes ungrammatical, as in (34b).

(34a) Ber-sin-nya saya sewaktu sedang mem-baca.
BER-sneeze-NYA 1SG when PROG AV-read
‘I sneeze while reading.’

(34b) *Ber-sin-nya sewaktu sedang mem-baca saya.
BER-sneeze-NYA when PROG AV-read 1SG
‘I sneeze while reading.’

This restriction in Indonesian may suggest that the verb and its subject must be adjacent, a feature not found in English presentationals. This could be necessitated by the suffix -nya; morphological considerations are discussed in Section 5.1, but remains a question for future research.

4 Analysis

Given all the similarities between Indonesian ber-V-nya constructions and the English presentational-there constructions, we might propose a similar analysis, an expletive account, for these novel sentences as has been proposed for their English equivalents. One clear difference is that the expletive analysis for ber-V-nya constructions would include a null expletive, whereas it is overt in English. These ber-V-nya constructions would then have the structure given in (35) below:

(35) [TP Exp T [sP [vP ber-V-nya DP] [XP ]]] (modified from Deal 2009:47)

This analysis accounts for the verb-initial ordering and the provided similarities to the English presentational-there constructions. One crucial prediction of this analysis, however, is that the postverbal subject would be located within the VP proper. There are three pieces of evidence against this analysis:

8 Thank you to an anonymous reviewer who suggested the inclusion of this section.
inability to co-occur with hanya, necessity of the postverbal subject, and the co-occurrence with definite subjects.

4.1 Inability to co-occur with hanya
Chung (2008) uses hanya ‘only’ as a diagnostic to determine the position of the subject in her verb-initial clauses. She argues that hanya associates with a focused constituent in its c-command domain. This predicts that a subject that is located within the VP can be focused through the use of hanya, but one located higher (such as in spec,TP) cannot. Chung provides the data in (36a-b) as evidence of this:

(36a) *[Wanita itu] hanya akan mem-baca buku ini.
Woman that only FUT AV-read book this
‘Only the woman will read this book.’

(36b) *[Buku ini] hanya saya baca.
Book this only 1SG read
‘I read only this book.’ (Chung 2008:27a-b)

This diagnostic can also be used here, to determine the position of the subject in ber-V-nya constructions. If the subject is located within the VP, it should be possible to use hanya to focus the postverbal subject. This does not seem to be the case, as exemplified by (37) below.

(37a) Ikan hiu hanya ber-enang di laut.
Shark only BER-swim in sea
‘Sharks only swim in the sea.’ but not ‘*Only sharks swim in the sea.’

(37b) Hanya ber-enang-nya ikan hiu di laut.
Only BER-swim-NYA shark in sea
‘Sharks only swim in the sea.’ but not ‘*Only sharks swim in the sea.’

In (37a), canonical SVO word order is given, and hanya cannot be used to focus the subject ikan hiu ‘shark’. This same pattern is found in (37b), where the ber-V-nya version is given. If the structure of this construction is (35), this would be unexpected, as the postverbal subject is located within the VP, and should be able to be focused by hanya. Instead, the subject must be outside the VP, seemingly in a position similar to the canonical SVO structure. Since hanya adjoins to the VP, only constituents below will be available for association with it. In the canonical SVO order in (37a), this can be the verb itself, but it cannot be the subject (which has moved into the spec,TP). In the ber-V-nya construction in (37b), these same facts hold: hanya can associate with the verb berenang, but cannot associate with the subject (see Figure 2 in Section 4.4 below for a structure of this construction). If these constructions had the pattern in (35), the subject ikan hiu would be within the VP. This suggests that such an analysis is untenable.

Interestingly, this behavior of ber-V-nya constructions contrasts with ada-constructions, which would be the closest kind of construction to ber-V-nya were this to be analyzed with the structure in (35), as exemplified by the ada-sentence given in (38) below:

(38) Hanya ada satu buku di atas meja.
Only there one book on table
‘There is only one book on the table.’

While there is some disagreement on what type of construction ada-constructions are (see Tjung 2005 for a discussion), they are likely the closest type of construction to an existential in Indonesian. It is therefore
interesting to note that ada-constructions do not share these same features of ber-V-nya constructions, which suggests that the subject is not located in the same position in sentences like (38).

4.2 Necessity of postverbal subject
If these ber-V-nya constructions have the structure given in (35), then we might expect that these constructions would allow omission of the postverbal subject, similar to the impersonal passive, found in German (Cardinaletti 1990, among others) and Icelandic (Bobaljik & Jonas 1996):

(39) Es wurde getanzt.
It was danced
‘There was dancing.’ (Cardinaletti 1990)

This is not possible in ber-V-nya constructions. The postverbal subject is always required:

(40a) Ber-sin-nya saya di luar karena alergi.
BER-sneeze-NYA 1SG outside because allergies
‘I sneeze outside because of allergies.’

(40b) Ber-sin-nya di luar karena alergi.9
BER-sneeze-NYA outside because allergies
‘He/she sneezes outside because of allergies.’ but not ‘*There is sneezing outside because of allergies.’

This is unexpected under a null expletive analysis as given in (35). Like with hanya above, this also contrasts with ada-constructions, which can be interpreted as more impersonal, as seen in (41):

(41) Ada yang ber-sin di luar.
There COMP BER-sneeze outside
‘There is someone sneezing outside.’

4.3 Co-occurrence with definite subjects
Indonesian ber-V-nya constructions allow a multitude of different subjects, including definite subjects and indefinite subjects. The data given below shows the subject as a pronoun (42), as definite10 (43), and as a quantified phrase (44).

(42a) Ber-sin-nya saya waktu sedang mem-baca.
BER-sneeze-NYA 1SG while PROG AV-read
‘I sneezed while reading.’

(42b) Ber-nyanyi-nya mereka di gereja.
BER-sing-NYA 3PL in church
‘They sing in church.’

(43a) Ber-main-nya anak itu dengan main-an.
BER-play-NYA child that with play-NOMI
‘The child plays with the toys.’

---

9 This is only grammatical under the interpretation of -nya as the third singular pronoun, not with -nya as being discussed in this paper.

10 There is some disagreement on that status of itu as a definite marker, which is beyond the scope of this paper.
This is also unexpected if the analysis given in (35) were correct, as we would expect to see the Definiteness Restriction (Milsark 1974; Jenkins 1975; Lasnik 1992, among others). It should also be noted, however, that these constructions also readily accept indefinite subjects.

In (45), anak ‘child’ is preceded by the classifier seorang. Nouns are always indefinite when co-occurring with a classifier in Indonesian. Postverbal subjects in ber-V-nya constructions can be both definite and indefinite. Again, this contrasts with ada-constructions, where the Definiteness Restriction does seem to be in effect:

4.4 Chung’s (2008) VP-fronting analysis
Given these facts, the null expletive analysis cannot be the correct analysis. An alternative analysis would be to consider other analyses of verb-initial constructions in Indonesian. One option is Chung’s (2008) analysis for verb-initial sentences in some dialects of Indonesian: the VP fronts to a position left of spec,TP (where the subject is located), to some functional head F. Chung proposes this to account for data like the following:

Such constructions are similar to ber-V-nya constructions in a few ways: one, the verb has fronted to a position higher than the subject, and two, the aspect marker sudah is preverbal. Of course, verb-initial sentences of this type are not identical to ber-V-nya constructions, as they do not necessitate the inclusion of the suffix -nya, and furthermore, are not intransitive. Nonetheless, it could be an attractive option to utilize an existing analysis for sentences of a similar type. Chung analyzes (47) as having the structure below:
Figure 1: Chung’s analysis of verb-initial constructions in Indonesian

Chung’s analysis of sentences such as (47) is to assume that the structure of the VP can differ depending on dialect, and in a dialect that allows the verb-initial order found in (47), the VP can be fronted to a position left of the subject in spec,TP. Importantly, the subject in SVO clauses is also located in this position (spec,TP) under Chung’s analysis. She assumes that subjects do not undergo any other movement within verb-initial clauses than they do in SVO clauses. Instead, she argues that the VP raises outside the clause proper, to the specifier of some functional head F. We might want to approach ber-V-nya constructions with the same approach.

This analysis accounts for the verb-initial ordering and has one crucial difference from the null expletive analysis: the postverbal subject in Chung’s analysis is not located within the VP, but in spec,TP, which would explain all the facts outlined in the sections above. Crucially, the location of the subject in spec,TP explains why subjects can be both definite or indefinite: an analysis that posits that the subject is located with a topic phrase, for instance, would not account for the possibility of indefinite subjects. And the analysis given in (35) cannot account for the possibility of definite subjects. Furthermore, the impossibility of association of focus with the subject in (37b), for instance, would be accounted for under this analysis. This is exemplified in Figure 2, where hanya adjoins to the fronted VP and therefore does not c-command the subject ikan hiu, but does c-command the VP:
5 Conclusion and lingering questions
This paper has presented what I have titled ber-V-nya constructions, novel data from Indonesian that represents not only a new function of the suffix -nya, but expands knowledge on potential functions of what has been previously dubbed the intransitive prefix ber-. It has also revealed the connection between these affixes and non-canonical word order in Indonesian, suggesting that morphology may play a role in triggering non-standard verb-initial order in some constructions. I have claimed that such constructions are a type of Indonesian presentational, analogous to presentational-there constructions in English, by providing evidence of many notable similarities between these two constructions. Lastly, I have provided two potential analyses to account for such data, but have presented evidence in favor of Chung’s VP-fronting analysis.

There are still many lingering questions I hope to answer in the future. One obvious question to ask at this juncture is what the nature of the interaction of these two morphemes are. Namely, is ber- + -nya a circumfix, that only functions in this capacity when the two interact? Or is this simply a function of -nya that occurs with intransitive verbs, some of which happened to be prefixed with ber-? To explore this further, it would be helpful to consider what happens when -nya is suffixed to bare intransitive verbs.

5.1 Morphological considerations
Suffixation of the type that was mentioned above is possible; this is obvious from Grangé’s (2015) subordinated NPs, many of which are intransitive verbs. However, I have already noted in Section 2 that the ber-V-nya constructions offered here do not fit into Grangé’s analysis. There do seem to be constructions of this type that, on the surface, appear to be identical to ber-V-nya constructions. Consider the constructions in (48).

(48a) Saya kembali ke Indonesia.
    1SG return to Indonesia
    ‘I return to Indonesia.’
Like *ber-V-nya* constructions, the sentence given in (48b) is necessarily verb-initial when -*nya* is attached to the verb *kembali* ‘back/return’. It also has the subject following it, and additional information in the form of a locative phrase. However, I offer up evidence that such data is not equivalent to *ber-V-nya* constructions, by showing some novel data of bare intransitive verbs being suffixed with -*nya* that differ crucially from the *ber-V-nya* constructions in two important ways: restriction on tense/aspect, and co-occurrence with negation. Both of these will be detailed below. Furthermore, it could be of import that all the constructions given in this section of ‘bare’-*V-nya* are judged to be significantly less marked than *ber-V-nya* constructions by speakers of Indonesian.

5.1.1 Restriction on tense/aspect

Bare intransitive verbs suffixed with -*nya* cannot co-occur with the progressive marker *sedang* nor the perfect aspect *sudah*, contra *ber-V-nya* constructions. This is evidenced by (49) below.

(49a) Sistem pemerintahan *sedang* goncang saat ini.
    System government PROG shake recently
    ‘The government is collapsing nowadays.’

(49b) Goncang-*nya* sistem pemerintahan saat ini.
    Shake-*NYA* system government recently
    ‘The government collapses nowadays.’

(49c) *Sedang goncang-*nya* sistem pemerintahan saat ini.
    PROG shake-*NYA* system government recently
    ‘The government is collapsing nowadays.’

In (49a), the use of *sedang* without -*nya* and verb fronting is completely acceptable. Affixation of -*nya* and subsequent fronting of the verb is also possible, as in (49b), but is not possible when *sedang* is also used, as seen in (49c). This is true not of just this bare verb, but others as well. The perfect marker *sudah* is also disallowed, as seen in (50) below.

(50a) Saya *sudah* kembali ke Indonesia.
    1SG PERF return to Indonesia
    ‘I have returned to Indonesia.’

(50b) *Sudah kembali-*nya* saya ke Indonesia.
    PERF return-*NYA* 1SG to Indonesia
    ‘I have returned to Indonesia.’

Such data suggests that these constructions are incompatible with aspectual markers, which differs from *ber-V-nya* constructions. This is one piece of evidence against categorizing these constructions with those prefixed with *ber-*.
5.1.2 Co-occurrence with negation

Another difference between these bare intransitive verbs and those prefixed with ber- concerns negation. I have already both in Section 2 and Section 3 noted that ber-V-nya constructions cannot take sentential negation with the negator tidak; any construction, then, that would be categorized as a presentational of this type would also resist sentential negation. This does not seem to be the case with bare intransitive verbs that are suffixed with -nya, as seen in (51) below.

(51a) Percaya-nya saya kepada Tuhan.
Believe-NYA 1SG to God
‘I believe in God.’

(51b) Tidak percaya-nya saya kepada Tuhan.
NEG believe-NYA 1SG to God
‘I don’t believe in God.’

As evidenced by (51), even when affixed with -nya the bare verb percaya ‘believe’ can be negated, with the negator tidak occurring preverbally. Sentential negation seems to be possible with all bare verbs affixed with -nya that occur in this construction. Consider kembali ‘back/return’ below:

(52) Tidak kembali-nya saya ke Indonesia.
NEG return-NYA 1SG to Indonesia
‘I didn’t return to Indonesia.’

These constructions’ ability to co-occur with negation is another difference between them and ber-V-nya forms. Again, this suggests that these are actually two different underlying constructions.

5.1.3 Circumfix or separate affixes?

The above findings suggest that ‘bare’-V-nya constructions are not the same as ber-V-nya constructions. What this may suggest, then, is that ber- and -nya are acting as a circumfix and must co-occur to derive this type of construction. This is not entirely far-fetched; Indonesian has other circumfixes, often ones that are made of two affixes that can also occur on their own (take ke--an, for example, where -an can also occur as a suffix without ke-). This is the case for ber-, which can occur not only as prefix on its own, but as part of the circumfix ber--an (Sneddon et al 2010). Analyzing it as a part of a circumfix with -nya, then, would not be implausible. Such an analysis could have implications for the morphology system of Indonesian and would require more fine-grained analysis. This remains a question for future research.

5.2 What functional head is the VP moving to?

Above, I argued that the VP in ber-V-nya constructions move to some functional head above the TP, following Chung (2008). Chung leaves the exact nature of this functional head a remaining question in her analysis, but is there any evidence from ber-V-nya constructions that might suggest what functional head this might be? One possibility is that this functional head is FocP (Rizzi 1997). Some evidence for this is the incompatibility of ber-V-nya to occur in questions. Consider the sets of canonical order questions and questions with ber-V-nya in (53) and (54) below.

(53a) Apakah mereka ber-kelahi men-guna-kan sepotong kayu?
Q 3PL BER-fight AV-use-KAN piece wood
‘Did they fight using sticks?’
5.3 Pragmatic considerations

Another obvious question concerns the pragmatic or discourse requirements that are necessary for these ber-V-nya constructions. It has been argued that there are specific discourse requirements for verb-initial constructions in Indonesian; specifically, that this non-canonical word order is used to highlight the predicate over the subject (Chung 1978; Kaswanti Purwo 1989). Both Chung and Tjung (2008) and Cole, Hermon, and Tjung (2005) have hypotheses about the discourse requirements of verb-initial constructions in Indonesian. Chung and Tjung (2008) have proposed a constraint, the Parallelism in VP-First Clauses (PV), which states that the subject of a verb-initial clause, if it is an external argument, must be aligned with the topic. This is a slightly scaled-down version of Cole, Hermon, and Tjung’s (2005) Parallelism Hypothesis (PH), which states that ‘the focus or new information must occur in the predicate and the subject must be the old information (topic)’ (562). If we are to assume that ber-V-nya constructions have the same syntactic analysis as other verb-initial constructions in Indonesian, we might want to also assume they have similar discourse requirements as well – so it would certainly be of interest to see if either the PV or the PH apply to the constructions discussed in this paper.

However, if, as I have argued above, these ber-V-nya constructions are analogous to the presentational-there constructions in English, we might expect that they would have similar discourse requirements as well. Ward and Birner (2001) argue that the post-verbal NP in the presentational-there constructions must be discourse-new but can be either hearer-new or hearer-old. It would be interesting to see if this requirement holds for ber-V-nya constructions. This, too, remains a question for further research.

References


UNTANGLING THE TAGALOG CLITIC CLUSTER

Russell Tanenbaum
Stony Brook University
russell.tanenbaum@stonybrook.edu

Abstract
This paper presents a syntactically-grounded account of the location and cluster-internal order of Tagalog second-position clitics, based on obligatory V-to-C head movement. Crucial to this analysis is the assumption that both the non-pronominal and pronominal clitics instantiate heads within the extended left periphery (c.f., Sportiche 1996, Cinque 1999). This approach derives in one fell swoop: the VSO word order of Tagalog, the fact that all Tagalog clitics occur in second-position, as well as the fact the non-pronominals appear in “mirror” order with respect to how they are presumably base-generated. Additionally, I examine data pertaining to the limited splitability of fronted nominals (in essence, constructions where “high” clitics can interpolate themselves and “low” clitics cannot) and take it as evidence that the proposed roll-up mechanism stops halfway through C. Empirical parallels to Slavic are highlighted, showing that the claims herein have wider applicability beyond Tagalog and Austronesian.

Keywords: Tagalog, syntax, clitics, particles
ISO 639-3 codes: tgl, trv, ces, rus

1 Introduction
For linguists interested in the behaviour of clitics (those tiny elements of a sentence whose prosodic weakness locates them at some midpoint between affixes and full-fledged words), Tagalog has long proved an interesting case study. Whereas some languages exhibit purely pronominal clitics (such as Romance, the starting point for most generative theories of clitics), and others purely non-pronominal ones (such the sentence-final enclitic particles of East and South-East Asia, or the second-position clitics of Russian), Tagalog is special in possessing a wide array of both pronominal and non-pronominal clitics, the latter encoding such varied functions as politeness, question-marking, evidentiality, and focus. Moreover, they all belong to a single cluster located unexceptionally in the second-position of the sentence, where a strict clitic order obtains. One particularly puzzling aspect of their order is the fact that both clitic type and syllable count have a role in where a given Tagalog clitic shows up in the cluster, as illustrated in the overview in (1):

(1) **Overview of the Tagalog clitic cluster (modified from Anderson 2008)**
   a. Monosyllabic pronominals always precede non-pronominals.
   b. Non-pronominals have an internal ordering among themselves.
   c. Non-pronominals always precede disyllabic pronominals.

Confronted with this data, Anderson (2005, 2008) argues that non-pronominal clitics are merged before pronominal clitics in the derivation of a sentence, by the same logic that inflectional morphemes are merged before derivational ones in word-formation. For him, OT-constraints LeftMost and NonInitial, derive second-position-ness in general and outlying monosyllabic pronominals are accounted for by the additional processes of “Stray Adjunction”. Kaufman (2010) builds on this OT-inspired approach, proposing an architecture of the Tagalog left periphery, and understanding the clitics as feature bundles that freely adjoin to any phrase where

---

1 I would like to extend my thanks to John Bailyn, Dan Finer, Francisco Ordoñez and Richard Larson for their support and guidance in this ongoing project, as well as the audience at AFLA25, especially Benjamin Brosig and Kie Zuraw, for help with new data. Lastly, a thank you to the anonymous reviewers for their valuable comments and suggestions for this paper. All its shortcomings remain my own.
they can be interpreted, modulo certain c-command requirements between pronominals and their associated argument positions. Though interesting and full of compelling arguments, these accounts nevertheless leave several questions unanswered. For instance: What makes the clitics form a cluster in the first place? And what determines the internal order of the cluster’s pronominal and non-pronominal domains?

In this paper, I propose a syntactically-grounded account to the puzzle of Tagalog clitics. I claim that observed clitic surface positions are a result of obligatory V-to-C rollup-style head movement. So long as we take both types of clitics to instantiate heads somewhere between V and the highest node of C, their clustering is an immediate consequence of this movement. This account additionally explains: Tagalog’s default verb-initial word order; the fact that among disyllabic pronominals, subject clitics always precede object clitics; and why the non-pronominal clitics appear in an order directly opposite the one that Cartography predicts. Section 2 provides a general overview of the Tagalog clitic facts. Section 3 reviews the analyses of pronominal clitics by Sportiche (1996) and Wurmbrand (2013), which provides the theoretical foundation for the present account. The analysis is presented in section 4, with step-by-step derivations. Section 5 deals with the so-called “phonological housekeeping” – the cases where the syntactic analysis falls short of deriving the observed surface order, and phonological processes must be appealed to to finish the job. The same approach is shown to successful explain analogous phenomena in Slavic. Section 6 addresses remaining problems. Section 7 summarizes and concludes.

2 General clitic ordering facts

The Tagalog clitic cluster occupies the second position in a clause. Furthermore, within this cluster, the non-pronominal clitics are sandwiched in-between the monosyllabic and disyllabic clitic pronouns. For instance, in the example in (2), modified from Billings and Konopasky (2002:4), the cluster consists of the monosyllabic second-person-singular object pronoun ka, followed by the interrogative particle ba, followed by the third-person-plural subject pronoun nila. The clitics encliticize to the verb, illustrating default verb-initial word order.

(2) Nakita ka ba niya?
PERF.see.ACC 2SG.OBJ.PTT Q 3SG.SUB
‘Did he see you?’

Of all of Tagalog’s 14 pronouns, only 3 are monosyllabic: ka, ko, and mo:

<table>
<thead>
<tr>
<th></th>
<th>Subject Clitics</th>
<th>Object Clitics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>ko</td>
<td>ako</td>
</tr>
<tr>
<td>1pl (incl)</td>
<td>natin</td>
<td>tayo</td>
</tr>
<tr>
<td>1pl (excl)</td>
<td>namin</td>
<td>kamí</td>
</tr>
<tr>
<td>2sg</td>
<td>mo</td>
<td>ka</td>
</tr>
<tr>
<td>2pl</td>
<td>ninyó</td>
<td>kayó</td>
</tr>
<tr>
<td>3sg</td>
<td>niyá</td>
<td>siyá</td>
</tr>
<tr>
<td>3pl</td>
<td>nilá</td>
<td>silá</td>
</tr>
</tbody>
</table>

2 Abbreviations used within the examples of this paper are as follows: ADD additive (focus), ADDR address, CONTEMP contemplative, DEIK deixis, EVID evidential, EXCL exclusive (focus), PERF perfective, PROG progressive, PRT particle, PTT promotion-to-trigger.

3 Glossing of (2) is recast under the terms of the analysis of Wurmbrand (2013) (c.f, footnote 5).
Taking a closer look at the non-pronominals, we notice, first of all, that there is no shortage of them. The table below (which is based on Anderson 2005, but updated to fit the current analysis) lists 16 of the most common ones. As suggested by their English glosses, these clitics fulfill a variety of functions, such as clause-typing, deixis/aspect, focus, and evidentiality.

**Table 2: Tagalog non-pronominal clitics (from Anderson 2005)**

<table>
<thead>
<tr>
<th>Clitic</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ba</td>
<td>(interrogative)</td>
</tr>
<tr>
<td>kasi</td>
<td>‘because’</td>
</tr>
<tr>
<td>kaya</td>
<td>(speculation)</td>
</tr>
<tr>
<td>daw</td>
<td>(reported speech)</td>
</tr>
<tr>
<td>din</td>
<td>‘too’</td>
</tr>
<tr>
<td>ho</td>
<td>(politeness)</td>
</tr>
<tr>
<td>lamang</td>
<td>‘only’</td>
</tr>
<tr>
<td>man</td>
<td>‘even’</td>
</tr>
<tr>
<td>muna</td>
<td>‘for a while’</td>
</tr>
<tr>
<td>na</td>
<td>‘already’</td>
</tr>
<tr>
<td>naman</td>
<td>‘instead’</td>
</tr>
<tr>
<td>nga</td>
<td>‘really’</td>
</tr>
<tr>
<td>pa</td>
<td>‘still’</td>
</tr>
<tr>
<td>pala</td>
<td>(surprise)</td>
</tr>
<tr>
<td>po</td>
<td>(politeness)</td>
</tr>
<tr>
<td>sana</td>
<td>(optative)</td>
</tr>
<tr>
<td>tuloy</td>
<td>‘as a result’</td>
</tr>
<tr>
<td>yata</td>
<td>(uncertainty)</td>
</tr>
</tbody>
</table>

Moreover, it is not uncommon to find four or five of them together in a single cluster, not including the pronominals clitics that would flank them in the absence of full DP arguments (note that pronouns in Tagalog never double their antecedents in a sentence, but rather are in complementary distribution with them). When that happens, a preferred ordering obtains, as represented in (3) and exemplified in (4) and (5). The examples data here have been adapted from Schachter and Otanes (1972:414), with glosses changed to fit the current proposed analysis.

(3) **Tagalog Particle Order**

na/pa ‘now/still’ > din ‘also’ > lang ‘only’ > daw ‘they say’ > po ‘sir’ > ba (Q)

Deik° ADD-Focus° EXCL-Focus° Evid° Addr° Force°

(4) **Nag-aaral na lang po ba si Juan ng Tagalog?**

Study Deik° excl-Focus° Addr° Force° PTT Juan DET Tagalog

‘Does Juan now just study Tagalog, sir?’

(5) **Uulan din daw ba bukas?**

CONTEMP-rain ADD-Focus° Evid° Force° tomorrow

‘Will it also apparently rain tomorrow?’

Interestingly, the preferred order of Tagalog non-pronominal clitics is the reverse of that of the base positions of the functional heads that these clitics are presumed to instantiate. The clitic *na*, which is often glossed as ‘now’ or ‘already’, can be understood as a realization of deixis, in that when it combines with an event-denoting predicate it gives rise to an interpretation of the event having transpired before some contextually-determined point of reference (much like English perfect tense). As such, it is clearly “inward-facing” from the standpoint of Rizzi (1997), and therefore should be at the low end of Tagalog’s left periphery.

---

As pointed out by a reviewer, the term *aspect* may more straightforwardly explain the function of *na* and *pa*. I employ the term *deixis* in to keep the analysis broad enough to capture the comparable the Chinese sentence-final particles *le* and *ne* à la Sybesma and Li’s (2007) mapping of the Chinese left periphery, though nothing major hinges on this terminological choice.
The interrogative clitic *ba*, by contrast, is “upward-facing”, realizing interrogative force, and therefore belongs at the high-end of the periphery. And in-between sits *lang* ‘only’, realizing (restrictive) Focus.

One final fact of Tagalog clitic ordering that deserves attention is that when sequences of disyllabic pronominal clitics appear in a sentence, such as (6), subject pronouns invariably precede object pronouns. This generalization in fact vacuously extends to all pronoun pairs in the language, given the paucity of monosyllabic pronouns to begin with, and the fact that the one viable candidate pair (*ko* + *ka*) is supplanted (*kita*).

(6)  
Nakita  
niya  
ako.  
PERF.see.ACC  3SG.SUB  1SG.OBJ.PTT  
‘He saw me.’

As we shall see in the next section, this last fact in particular begs a syntactic explanation5.

3  A Word on Wurmbrand’s (2013) Sportichean analysis of Tagalog pronominals
In a 2013 manuscript that explores multiple aspects of the Tagalog verbal and clausal domain, Wurmbrand presents an analysis of Tagalog pronominal clitics based on Sportiche’s (1996) null pronoun-based account of Romance pre-verbal clitics. Specifically, she posits a clitic projection between CP and TP, calling it CP$_{CL}$. Into this projection’s specifier will raise a pro-CL, i.e., a null DP, that has been generated in the appropriate argument position. The clitic is then the spellout of the head of CP$_{CL}$. The pro-CL binds as an operator/quantifier, while the clitic supplies agreement features. It is assumed that in the case of clitic doubling languages, unlike Tagalog, the pro-CL is swapped for an overt DP. The mechanism is schematized in (7).

(7)

Elsewhere, to avoid locality violations, Wurmbrand relies on tucking-in to derive vP structures wherein the verb agrees with the object (what she glosses as ACC on the verb). In the case of multiple pronominal second-position clitics, however, tucking-in is of no help to her, as she is committed to the (in my opinion, plausible) assumption that every clitic is in fact a functional head, and therefore must correspond to a different projection. Consequently, she proposes the derivation shown in (8). This is to be understood as an analysis of sentences such as (6), where the subject clitic precedes the object clitic6.

5  As pointed out by an anonymous reviewer, this order of subject preceding object—a phenomenon more commonly referred to by other terms, e.g. *actor* preceding *theme* (Billings 2005)—is in fact a property of phrasal arguments as well. How to integrate this fact into the present account is not yet clear.

6  Note that Wurmbrand differs from other Austronesianists in defining *subject* as ‘external argument’ and *object* as ‘internal argument’, rather than according to which argument is ang marked. Ang-marking, for its part, is referred to as *Promotion To Trigger* (PTT), and is distinct from case. I follow her terminology.
In Wurmbrand’s account, first the object pro-CL moves to [spec, Cl1] (green arrow), with the head spelling out as ako. Then the subject pro-CL moves to [spec, Cl2] (pink arrow), with the head spelling out niya. A complicated system of star-marked interveners (based on Bošković 2011) is appealed to to justify the crossing of paths in (8), since otherwise we’d expect the subject pro-CL to be the first one that moves. To make her derivation work, Wurmbrand is required to say that the first (green) movement is A-movement, while the second (pink) movement is A’. While this distinction does play a role elsewhere in Wurmbrand’s larger analysis, it still seems arbitrary and stipulative that within a single cluster, one clitic’s specifier should be A, and the other’s A’. Even in her own system it creates a lookahead problem which requires a lot of fancy footwork on her part to circumvent. In the next section, I propose an alternative account, which builds off of Wurmbrand’s but avoids the problems she runs into.

4 Proposal: Clitic clusters as a consequence of V-v-T-C rollup
Under the assumption, contra Wurmbrand (2013), that no A/A’ distinction exists between the different clitic positions in (8), we would expect the subject pro-CL to be the first to move at the moment the lower clitic head is joined. With the subject pro-CL’s featural requirements (related to phi-licensing) discharged, the object pro-CL then becomes the one and only candidate for movement into the spec of the higher clitic head. This is schematized in (9).
(9) Raising first of subject, then object, pro-CL (contra Wurmbrand (2013))

Of course, this problematically derives the incorrect pronominal clitic order *ako* > *niya*. All is redeemed, however, if we assume that the structure in (9) is simply the input to cyclic successive head movement, whereby V raises and left-joins first to v, then to T, CL1, Cl2, and (expanded) C, respectively, as shown in (10).

(10) Head-movement and left adjunction, yielding V+v+T+Cl1+Cl2+C:
This would capture two additional phenomena, apart from the subject > object clitic order. First, it would provide a purely syntactic explanation for why Tagalog syntax is by default verb-initial. (Whereas it’s unclear exactly how the second-position-ness of second-position effects are arrived at in Wurmbrand’s system.) Second, it would straightforwardly explain why the order of non-pronominal clitics is precisely the inverse of that of their presumed base-merge positions.

By hypotheses, the non-pronominal clitics in a sentence like (4), repeated below as (11), instantiate separate heads within an expanded CP. Drawing from work on the Cantonese clausal domain by Sybesma and Li (2007:1778), I understand Tagalog na, at the lowest end, as the head of DeikP. This functional head, explains Sybesma and Li, “establishes a link with the speech moment,” and thereby “anchors the sentence to the time axis of the real world.” Deik here stands for “deictic,” due to na’s reference-linking function. Above that is FocusP, headed by lang which denotes exclusive focus – glossable as only – and corresponding to Cantonese ze⁸. One step up, following Slocum’s (2016) work on the syntax of vocatives, I assume the existence of an AddrP (as in address), generated above Focus and below Force, that serves to mediate the relationship between speaker and hearer. In Tagalog, this head has three possibilities for phonological realization: po, ho, and Ø. Generally speaking, po is used when addressing elders and superiors in formal settings, with ho as a somewhat less formal variant thereof. (I therefore loosely gloss these addressee-related elements as ‘sir.’) Finally, at the high end of the CP domain, is ForceP, headed by ba, whose function is obviously one of clause typing, as its addition to a sentence changes statements into questions. All together, this yields the structure in (12a), corresponding to the relevant portion of the sentence in (11). As a result of the head-movement schematized above in (10), however, the nonpronominal clitics in (12a) are reordered in the narrow syntax, deriving their observed surface order in (12b).

(11) ((=4)) 
Nag-aaral na lang po ba si Juan ng Tagalog?
study DEIK FOCUS ADDR FORCE PTT Juan DET Tagalog
'Does Juan now just study Tagalog, sir?'

(12)

a. ForceP
   ba
   AddrP
   po
   FocusP
   lang
   DeikP
   na
   .......

b. ForceP
   Force
   .......
   Addr
   po
   Focus
   na

(12a) corresponds to (11), and (12b) to (12a).

---

⁸ While Cantonese has no apparent overt realization of Addr°, it does realize Deik° (la), Focus° (ze) and Force° (ma), all of which correspond both in meaning and surface order to their Tagalog counterparts, suggesting a true universality of clitic ordering for the sequence Deik > Foc > (Addr >) Force:

(i) la (Deik) > ze (Focus)
   Ngo jao dai yat nin la ze.
   I more big one year T/deik only
   ‘I’ve merely gotten one year older.’

(ii) ze (Focus) > ma (Force)
   Keoi faan jat jat ze ma?
   He return one day only Q
   ‘Is he only coming back for one day?’
Next, by replacing *si Juan* ‘Juan’ with *ka* ‘you’, as in (13), we are able to see the interaction of pronominal and non-pronominal clitics. The attested word order is correctly predicted by the expanded analysis in (14).

(13)  
\[ \text{Nag-aaral ka lang ho ba ng Tagalog?} \]

study 2SG.PTT FOCUS ADDR FOCUS DET Tagalog

‘Do you only study Tagalog, sir?’

(14)

Not every sentence in Tagalog is verb-initial, however. Wh-words and negatives can also appear in initial position, separately or together, as (15-17) show (from Schachter & Otanes 1972):

(15)  
\[ \text{Hindi nagpapahinga ang bata.} \]

NEG rest.PROG PTT child

‘The child is not resting.’

(16)  
\[ \text{Bakit nagpapahinga ang bata?} \]

Why rest.PROG PTT child

‘Why is the child resting?’

---

9 This analysis echoes research elsewhere in Austronesian syntax; in particular, Rackowski’s (1998) work on Malagasay adverbs, where she concludes, based on evidence such as the sentences in (i), that Malagasay post-verbal adverbs appear in the reverse order of Cinque’s (1999) hierarchy (provided in (ii)) due to successive cyclic movement.

a. Tsy *tena mbola* mahay mandihay Rakoto

‘Rakoto still really doesn’t know how to dance.’

b. Manasa lamba *tsara foana* Rakoto.

Wash clothes *well always* Rakoto.

‘Rakoto always washes clothes well.’

solitamente > mica > gia > piu > sempre > completamente > tutto > bene
usually > not > already > any longer > always > completely > all > well

---

87
When clitics are added to (15) and (16), they invariably appear in second position, immediately after the negative hindi or wh-word bakit. This is where we would expect to find them anyway, seeing as on the present analysis clitics are “indigenous” to the left periphery. Sentences with both negatives and wh-words, like (17), on the other hand, have the potential to diagnose with more accuracy just where in the left periphery the clitics are located. As it turns out, any non-pronominal clitic may appear after the string bakit hindi, but only certain ones may appear in the middle of it, as seen in (18-19) (modified from Schachter & Otanes 1972:432-433):

(18) Bakit hindi ba/po/na/pa nagpapahinga ang bata?
    Why NEG Q/sir/now/still rest.PROG PTT child
    ‘Why is the child (?/sir/now/still) not resting?’

(19) Bakit ba/po/*na/*pa hindi napapahinga ang bata?
    Why Q/sir/now/still NEG rest.PROG PTT child
    ‘Why is the child (?/sir/*now/*still) not resting?’

Putting aside the question of whether bakit and hindi are to be analyzed as adjuncts or functional heads in their own right, if we assume that bakit occupies a position at the high-edge of the C-domain, while hindi occupies a position toward the middle of it, then the restrictions on which clitics can intervene are not at all surprising. Na and pa, both potential realizations of Deik, are argued above to be at the bottom of the C domain, and therefore are prevented from preceding negation due to C-selection10.

One empirical hurdle to this analysis is its prediction that all Tagalog pronominals should precede non-pronominals in clusters, whereas in fact the monosyllabic ones precede them, and the disyllabic ones follow them. Recall the generalization in (1), rephrased below as (20).

(20) Tagalog clitic cluster (broad view)
    monosyllabic pronominals > non-pronominals > disyllabic pronominals

A modicum of phonological realignment is therefore still needed to keep the theory in line with the facts. But if, as claimed, the cluster-finality of disyllabic pronouns is merely a post-syntactic heavy shift effect, then we would expect that any monosyllabic variants of these normally disyllabic pronouns should be able to appear before the non-pronominals, i.e., in the position that narrow syntax has put them in post-head-movement. The only pronouns that admit such phonetic variance are second-person plural ninyo (~n’yo) and third-person singular niya (~n’ya). Happily, as shown in (21) with data from Billings (2005:313), both of these orders, though a bit “strange,” are nevertheless “acceptable” – evidence that syntactic ordering is indeed sub.CL° > obj.CL° > non-pronominals.

---

10 The fact that ba, po and other high clitics are able to appear on either side of negation, not just before it as predicted, may indicate the existence of multiple NEG positions within the same CP domain, as suggested by Kaufman (2010).
5 The Phonological housekeeping

As cluster-internal ordering is sensitive to some extent to phonological weight, perhaps out of obedience to an OT-style constraint like *Heavy Last*, the next logical question to ask is just how big a role phonological constraints play in determining word order. According to Anderson (2005), their role is quite large indeed. He presents an analysis of Tagalog that attempts to account for second position clitics and morphological infixation in one fell swoop. In brief, both clitics and affixes in Anderson’s system are at the mercy of well-formedness constraints specific to either the word or phrase level. *LeftMost* requires that the clitic or affix appear at the left edge of the domain. *NonInitial* requires that it not be the first element. Together, these yield second position effects. Mitigating against them is the Faithfulness constraint *Integrity* – again, specific to either the word or phrase level – which when ranked high enough prevents clitics from appearing internal to XPs, and when ranked low enough permits even infixation, as shown in (22).

(22) Tagalog Infixation per Anderson 2005 (LeftMost, NonInitial >> Integrity[word])

<table>
<thead>
<tr>
<th>in</th>
<th>basa</th>
<th>b&lt;in&gt;asa</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERF</td>
<td>read</td>
<td>have.read</td>
</tr>
</tbody>
</table>

For Anderson, these constraints are assumed to be active on all levels of grammar. Generalizing them to syntax, he lays out what amounts to an OT account of Tagalog V1 word order (extending it in turn to V2 phenomena in Icelandic and other languages). In my account, on the other hand, the fact that Tagalog clitics are always in second position and Tagalog verbs generally appear adjacent to those clitics (either in first or third position) is to be understood as a consequence of the verb and the clitics being co-participants in the same syntactic process of head raising. OT constraints alone are, after all, ill-equipped to explain the “mirror order” of pronominal and non-pronominal clitics. But that said, a close examination of the data reveals that certain clitics – in particular, interrogative *ba*, subordinators *kasi* ‘because’, *man* ‘even though’, and address related *po/ho* – often occur farther to the left than a pure syntactic analysis (or even one enriched by cluster-internal prosodic inversion) predicts that they should. This is shown for *ba* in (23–25) (source: Schachter & Otanes, 1972:432).

(23) WHY + ba + NOT + verb (A modified version of (19))

<table>
<thead>
<tr>
<th>Bakit</th>
<th>ba</th>
<th>hindi</th>
<th>nagpahinga</th>
<th>ang</th>
<th>bata?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why</td>
<td>Q</td>
<td>NEG</td>
<td>rest.PROG</td>
<td>PTT</td>
<td>child</td>
</tr>
</tbody>
</table>

‘Why is the child not resting?’

(24) Adv + ba + Top + verb

<table>
<thead>
<tr>
<th>Bukas</th>
<th>ba</th>
<th>‘y.</th>
<th>aalis</th>
<th>siya?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomorrow</td>
<td>Q</td>
<td>TOP</td>
<td>FUT.leave</td>
<td>he</td>
</tr>
</tbody>
</table>

‘Tomorrow will he leave?’
(25) Fronted adverbial, split by ba

\[
\begin{array}{ccccccc}
\text{tomorrow} & Q & \text{GEN} & \text{night,} & \text{dance} & \text{SUB.PTT.they} & \text{DET} & \text{Fandango} \\
\end{array}
\]

‘Will they dance the Fandango tomorrow night?’

In (23), interrogative \(ba\) is separated from the verb by \(hindi\) ‘not’; in (24), it is the topic marker \(ay\) (here shortened to ‘y’) that separates \(ba\) from the verb; in (25), it is the second half of the fronted adverbial \(bukas ng gabi\) ‘tomorrow night’ that intervenes. In cases of fronted adverbials that are not separated from the rest of the sentence by the topic marker (\(ay\), one gets the impression that the fronted adverbial does what is otherwise the verb’s job of dragging up all the clitics to the front of the sentence, as in (26). The same can be said of the negative in (27).

(26) \[\begin{array}{cccc}
\text{tomorrow} & Q & \text{he} & \text{FUT.leave} \\
\end{array}\]

‘Tomorrow will he leave?’ (Schachter & Otanes, 1972:189)

(27) \[\begin{array}{cccc}
\text{NEG} & Q & \text{he} & \text{FUT.leave} \\
\end{array}\]

‘Will he not leave?’ (Schachter & Otanes, 1972:189)

A possible explanation for (27) might be that \(hindi\) is in fact a negative auxiliary/copula, successively-cyclically raising through the functional domain, while the verb \(aalis\) ‘leave’ remains in-situ. All clitics would thus be exactly where they belong. Indeed, other Tagalog negatives, such as \(wala\) ‘there is no…’ and \(huwag\) ‘don’t…’, participate in the same clitic-attracting behavior as \(hindi\), and differently from negatives in many other languages (Romance, Slavic, Finnish, Chinese…), negatives in Tagalog show no difference in form with respect to sentential negation versus fragment response (i.e., \(not = no\)). But one presumably cannot make the same case for the “secretly verbal status” of every non-verb in sentence-initial position. \(Bukas\) ‘tomorrow’, in (26), for example, is unlikely to be anything but a mere time adverb, and therefore is no candidate for raising to \(C^o\).

A second potential pure syntax solution, strong enough to cover both (26) and (27), might be to say that the verb, \(aalis\), does indeed raise to \(ba\) in the narrow syntax, but ultimately it’s the lower copy that’s pronounced. This would be in the same spirit as Bošković’s (2002) application of Chomsky’s (1993) copy theory of movement to a variety of syntax-phonology mismatches, including the fact that Serbo-Croatian auxiliary \(je\) is always last in the clitic cluster despite ellipsis evidence that points to the contrary.

(28) \[\begin{array}{cccc}
\text{she} & \text{him.DAT} & \text{him.ACC is introduced} & \text{and also he is him.DAT him.ACC introduced} \\
\end{array}\]

‘She introduced him to him and he did too.’

Another example from Bošković and Nunes (2007) that deserves mention due to its striking similarity to the Tagalog facts involves Northern Norwegian V2 and is given below in (29). Unlike the rest of Scandinavian, Norwegian apparently shows sensitivity to whether or not the element in first position is a full foot (i.e., two syllables long).

(29) a. \[\begin{array}{cccc}
\text{Korsen} & \text{kom} & \text{ho hit?} \\
\text{how came she here} \\
\text{‘How did she get here?’} \\
\end{array}\]

b. \[\begin{array}{cccc}
\text{Kor} & \text{du} & \text{kom fra?} \\
\text{where you came from} \\
\text{‘Where did you come from?’} \\
\end{array}\]

\[90\]
For Bošković & Nunes, in (29b) the verb does raise to C, it’s just the lower copy that gets pronounced:

(30)  *Kor kom, du kom, fra?

An analysis of (26) along the lines of (31) is therefore very tempting¹¹:

(31)  *Bukas aalis, ba siya aalis,?

However, neither of these explanations is capable of accounting for (24) and (25), since if a copy of the verb had moved up to the highest head of the left periphery, by all assumptions it would have carried the pronominal clitic with it, and all the clitics would be stranded in a cluster together, contrary to fact. Rather, it seems more likely that ba is separable from the rest of the cluster simply because the verb raises only so high within the expanded CP domain.

One piece of positive evidence in favor of this approach is the fact, mentioned above in passing, that the clitics kasi ‘because’, man ‘even though’, and ba all pattern together in that they obligatorily “constitute an immediate part of the initial component of a sentence or clause” forming a natural class (Schachter & Otanes 1972:440; Anderson 2008). Indeed, in Pearson’s (1998) analysis of Malagasy, he posits a Sub°, right above Force°, as the location for clause-initial subordinators. If this is true of Tagalog, such that kasi and man instantiate Sub°, and ba instantiates Force°, then it is no surprise that they are similarly distributed. (32) and (33) (from Schachter & Otanes 1972:348), demonstrate this fact.

Awakened PTT child. Noisy because you, PRT
‘The baby awakened. (That’s) because you were noisy.’

(33)  Mahal man ang sapatos, mabili ko rin.
Expensive even.though PTT shoes, bought I also
‘Even though the shoes were expensive, I bought them.’

Another argument involves the similarly exceptional behavior of ba’s Slavic counterpart, li – a functional element which happens to have identical cognates in Russian, Czech, Bulgarian and Macedonian (among others). In Czech, li is exceptional in that it can split a DP, whereas other clitics cannot. In (34), for example, the pronominal clitic mi cannot split the DP ten basnik ‘that poet’. In (35), however, the fronted DP lasce sve ‘one’s (own) love’ can be split by li, but not by the reflexive clitic se, which must instead appear to the right of it.

(34)  [Ten basnik] mi cte ze sve knihy.
That poet me reads from his book
‘That poet is reading from his book to me.’
(35) Fronted nominal, split by *li* (from Avgustonova & Oliva 1995)

Lásce-*li* své se v žítí budeš protiviti,
love.ACC-Q self’s.ACC refl.ACC in life fut.2SG oppose.INF,
žebrákem půjdeš světem.
beggar.INST go.2SG world.INST
‘If you oppose your love in your life, you will go through the world as a beggar.’

The superficial resemblance of (35) to (25), repeated below as (36) is striking.

(36) (=(25)) Fronted adverbial, split by *ba*

*Bukas* ba ng gabi, sasayaw *sila* ng pandango?
tomorrow Q GEN night, dance NOM.they ERG fandango
‘Will they dance the Fandango tomorrow night?’

Russian *li* can also split DPs. Interestingly, if the DP is long enough, it obligatorily splits it, rather than appearing phrase-finally, as shown in the example in (37) from Franks and King (2000:355).

(37) [Na ètom] *li* zavode on rabotaet?
In this Q factory he works
‘Does he work at this factory?’
(Cf. *Na ètom zavode li on rabotaet?)

For Franks and King, the alternative in (37) precludes an analysis where *na etom zavode* ‘in this factory’ is in the specifier of *li*. Instead, they conclude that fronted nominal sits in the specifier of *li*’s complement, the FocusP, and that prosodic inversion derives the observed surface order, where *li* surfaces after the first prosodic word. Therefore, (38) corresponds to the relevant portion of (37).

---

12 Frank and King (2000:351) cite asymmetries between Bulgarian and Macedonian *li* attachment that further support their claim that *li*’s surface position is subject to prosodic inversion. *Li*’s varied placement in (i–iv) is predictable if one makes the following three assumptions: One, *ti* and *go* are proclitics, but only in Macedonian. Two, *ne* causes the syllable after it to take stress, but only in Bulgarian. Three, *li* immediately follows the prosodic word to its right.

(i) Ne *ti* go dade *li*? [√Mac/*Bug]

NEG you.DAT it.ACC gave Q
‘Didn’t he give it to you?’

(ii) *Ne* ti li go dade [√Mac/*Bug]

(iii) *Ti* go dade li? [√Mac/*Bug]

you.DAT it.ACC gave Q
‘Did he give it to you?’

(iv) Dade li ti go? [√Mac/*Bug]
Finally, Holmer’s (2005) analysis of sentence final particles in Seediq indicates that in that language too, the head that carries the interrogative feature (pronounced ye) behaves differently from the other functional heads of the expanded C domain. While the aspectuals, evidentials, and discourse markers all surface sentence-finally, ye appears sentence-initially.

(39) Wada qyuxun alang Tongan sa.
   PAST rain village Tongan EVID
   ‘Apparently, it rained in Tongan village.’

(40) Ye su m-n-imah sino ciga?
    Q you drink wine yesterday
    ‘Did you drink wine yesterday?’

There is therefore ample crosslinguistic support that functional heads at the very edge of the left periphery, though at times superficially part of the same clitic cluster, are still subject to a different distribution under the right circumstances. Drawing on Franks and King’s (2000) analysis of Russian li, I conclude that the Tagalog verb raises in all cases to Foc° and no higher, with Foc° either null or spelled out as lang ‘only’ should its flavor be that of restrictive focus. Further, I posit that just above FocP is an optional TopP, phonetically realized either as ay or as comma intonation. The expanded CP domain is therefore as in (41):

(41) [ForceP ba [AddrP po/hO [TopP ay [FocP ø/lang [DeikP na/pa […]]]]]

A set of sentences is provided in (42), drawn from Schachter and Otanes (1972), with the corresponding syntactic outputs (after head-movement but before prosodic optimization by the grammar).

(42) a. Aalis ba siya? ‘Will he leave?’
    [ForceP ba [FocP aalis+siya]]
 b. Aalis ka ba? ‘Will you leave?’
    [ForceP ba [FocP aalis+ka]]
 c. Bukas ba, aalis siya? ‘Tomorrow, will he leave?’
    [ForceP ba [TopP bukas “,” [FocP aalis+siya]]]
 d. Bukas ba ng gabi, aalis siya? Tomorrow night, will he leave?
    [ForceP ba [TopP bukas ng gabi “,” [FocP aalis+siya]]]
 e. ?Bukas ng gabi ba, aalis siya? ‘Tomorrow night, will he leave?’
    [ForceP bukas ng gabi ba [TopP “,” [FocP aalis+siya]]]

Potential revisions to (41) include: reversing the order of Addr and Force (since po/hO tends to precede ba, and my analysis understands them - and indeed everything under ay - as not participating in the functional head rollup; and adding a possible Evid below the high Topic, for the same reason (i.e., the prevalence of “daw po ba” sequences).
The splitting of *bukas ng gabi* ‘tomorrow night’ by *ba* in the Tagalog sentence in (42d) ((= (25)) is therefore understood as structurally equivalent to the splitting of *na etom zavode* ‘in this factory’ by *li* in the Russian sentence in (37). Further support for this unified analysis comes from the fact that neither sentence incurs a focus interpretation of the pre-clitic element. This is expected if in each case the constituent remains unsplit in the narrow syntax.

6 Remaining Problems and Possibilities

As pointed out by an anonymous reviewer, the V-to-C head analysis presented in this paper leaves a portion of Tagalog data unaccounted for. For the sentences discussed in (42), the syntactic mechanisms proposed herein derive at least the correct adjacency of elements, even if recourse to post-syntactic realignments (such as prosodic inversion) are still needed. However, examples like (43a) from Kaufmann (2010:267), where the pronominal clitic *ako* is encliticized to the first, rather than the second, of the two preverbal elements, are particularly troublesome for a verb-raising analysis such as this one, which predicts negation should not be able to intervene between verb and clitic, as the two ought to be joined via head-adjunction:

(43) a. Madalas *ako* hindi makatulog.
    Often 1.SUB.PTT NEG NOM.sleep
    ‘Often, I don’t sleep.’

(43) b. Madalas *hindi* *ako* makatulog.
    often NEG 1.SUB.PTT NOM.sleep
    ‘Often, I don’t sleep.’

This raises the question of whether the data would be better explained under an approach that ascribes separate explanations to Tagalog’s 2P effects and its verb-first tendency. Other attempts at applying a head-movement approach to second position clitic phenomena include Legate’s (2008) account of Walpiri second-position clitics. As in Tagalog, verb-initial sentences are common in Walpiri, but so are adverb-initial and XP-initial sentences, with the clitics as a rule coming second. Legate makes the argument that verb-initial Walpiri sentences involve long-head movement of a verb into the attracting head of a functional projection high in the left periphery, whereas XP-initial sentences involve phrasal movement into that same high head’s specifier. Additionally, to explain how clitics are sometimes able to intervene between verbs and the so-called preverbs that precede them, she posits a mechanism (AFFIX) to invert clitics that are syntactically in first position with the following prosodic word. Legate’s data are compelling, but the theoretical cost is the need for is multiple, independent paths to V1. She further acknowledges (her footnote 35) that since the clitics themselves appear in the order C > Aspect > AgrP, then their clusterhood must either be a case of right-adjunction (contra antisymmetry assumptions) or clusterhood-by-adjacency, neither of which seems fully satisfying.

Another approach that may shed light on this puzzle is the one taken by Rackowski and Travis (2000) in analysing Malagasy and Niuean. Under their analysis, the VOS word order of Malagasy, and the mirror order of post-verbal adverbs, is explained via iterative VP movement through a series of immediately higher projections (headed by adverbs), followed by TP raising. Subsequently, the VSO word order of Niuean is explained via essentially the same movements, except with obligatory object shift to an AgrOP position that (unlike in Malagasy) is above TP in Niuean. Could this be the source of Tagalog clitics’ mirror order? Holmer (2005) dismisses the idea of an iterative VP movement analysis of the Tagalog 2P clitic cluster, in part because Tagalog already has head movement at its disposal (c.f., Aldridge’s (2004) claim of V-to-T raising) and that to him seems the simplest solution. Still, so long as sufficiently high AgrO and AgrS positions are posited via which to vacate the TP prior to the latter undergoing iterative raising, it’s not infeasible. It just becomes a question of which way the facts push us – a question which is still admittedly quite open.
7 Conclusion
It is only right that syntax be the first recourse when tackling seemingly intractable problems of word order, since word order is the raison d’être of syntactic inquiry. In that light, this paper seeks an answer to the question, How far can a syntactic explanation of Tagalog second-position clitics take us? Borrowing an idea from Sportiche (1996) and Wurmbrand (2013), I have assumed Tagalog pronominal clitics to be functional heads between C and T, into whose specifiers have moved proDPs that originate in canonical argument positions. The resultant structure is then the input for obligatory rollup-style V-to-C head movement, deriving the surface order. Under this analysis, a number of disparate phenomena (non-pronominals appearing in “mirror order”; pronominal obeying subject > object; verb-initialness of sentences) receive a unified explanation. Other facts, like the observation that certain clitics related to clause-typing, evidentiality, and politeness are “in the cluster but not of it”, or that fronted adverbials can also host clitics, strongly suggest that some elements of clitic placement are indeed of a post-syntactic nature. The analysis further predicts that, when surveying languages that have sets of non-pronominal clitics analogous to those of Tagalog, the typology of clitic orders should be constrained by the logical possibilities of how fully V raises to C, assuming a universal hierarchy of functional projections. How that prediction fares is a question for future research.

References
Bošković, Željko. 2011. Rescue by PF deletion, traces as (non)interveners, and the that-trace effect. Linguistic Inquiry 42.1:1-44.


Abstract
This study examines constructions of noun incorporation (NI) in Northern Paiwan and their morphosyntactic properties from both typological and theoretical perspectives. Noun incorporation in Northern Paiwan displays unique characteristics that are rarely seen in most NI languages. Incorporated nouns (INs) are restricted to peripheral arguments with respect to space, and case-marked noun phrases (NPs) extended with modifiers also incorporate. These special properties pose challenges for both lexical and syntactic approaches, neither of which alone can account for the phenomena. This study proposes a non-unified approach to NI in Northern Paiwan. It has been argued that Northern Paiwan NI involves two separate processes: lexical NI, which is morphologically derived, and syntactic NI, in which the IN surfaces as an oblique case-marked NP/DP (determiner phrase). The selection between two types of Northern Paiwan NI hinges on the constituency of the INs. These findings can shed light on the understanding of NI cross-linguistically.

Keywords: noun incorporation; Northern Paiwan; incorporated noun; idiosyncrasy; modifier stranding
ISO 639-3 codes: pwn; nci; tix; gup; niu; ckt

1 Introduction
Noun incorporation (NI) is widespread throughout the world, though it varies in different languages, and it is of considerable theoretical and typological interest (Bybee 1985, Baker et al. 2005). From a typological view, NI plays an important role in determining linguistic typology. NI is a salient characteristic of polysynthetic languages, whereas it is not observed in most non-polysynthetic languages (Baker 1996). From a theoretical perspective, NI bears on issues concerning the relationship between morphology and syntax (Mithun 1984, Baker 1988, Rosen 1989, Baker et al. 2005, Barrie & Mathieu 2016). NI represents a choice between syntactic and morphological techniques—which approach is more appropriate to express the semantic relationship between incorporated nouns (INs) and verbs? It also calls into question whether morphology and syntax are distinct components of the grammar. For the past few decades, a majority of research has focused on NI languages whereby the incorporated nominals are bare nouns (N₀) in NI constructions. Traditionally, NI has been considered a process in which nouns combine with verbs to produce a complex verb. In the process, the IN is expressed as a morphological root integrated into the verb rather than as an independent noun phrase (Sapir 1911, Bybee 1985, Rosen 1989, Baker et al. 2005). In recent years, however, languages that exhibit extended IN structures, such as noun phrases, have also been observed (Massam 2001).

This study examines the phenomena of noun incorporation in Northern Paiwan, an Austronesian language spoken in Southern Taiwan. This paper will demonstrate that Northern Paiwan NI covers a broad spectrum, ranging from regular NI to unusual case phrase incorporation. In Type I, a bare noun combines with
the host verb, giving rise to a verb-noun (V-N) complex verb,\(^3\) which can attract pronominal clitics, as shown in (1a). By contrast, in Type II, a nominal element, as shown in (1b), or a modified noun phrase, as shown in (1c), together with its case marker can all be incorporated into the verb, yielding an agglutinating-style scenario:

(1a) \[Na-s<em>a-pana=aken.\]
\[\text{go.to}<\text{INTR}^{4}>\text{-riverbank}=1\text{S.ABS}\]
‘I went to riverbank.’

(1b) \[K<em>si-tjay-palang=aken a mangtjez.\]
\[\text{be.from} <\text{INTR}>\text{-OBL-Palang}=1\text{S.ABS} \quad \text{LNK} \quad \text{come.back.INTR}\]
‘I came back from Palang’s (place).’

(1c) \[I-tua-tapaw-ni-camak=aken.\]
\[\text{be.in.INTR-OBL-hut-GEN-Camak}=1\text{S.ABS}\]
‘I am in Camak’s hut.’

The two types of NI shown in (1a-c) pose problems for both the lexical approach, in which NI is treated as a compounding process (Mithun 1984, Rosen 1989), and the syntactic approach, in which NI is analyzed as head-to-head movement (Baker 1988). Neither the lexical nor the syntactic approach alone is adequate in accounting for the situation in Northern Paiwan. In the light of the facts described above, the following research questions were proposed:

(i) What are the typological characteristics and properties of noun incorporation in Northern Paiwan?
(ii) Are the two types of Northern Paiwan NI syntactically or morphologically derived? If the two processes are derived separately, how is Type I NI differentiated from Type II in morphosyntactic characteristics and behaviors?
(iii) Are there rules or principles governing the selection of NI involved in spatial verbs? If so, under what conditions can the presence or absence of the oblique case markers \(tjay/tua\) be determined?

This paper will argue that the two types of Northern Paiwan NI involve non-unified grammatical processes. Type I NI is identified as lexical NI and Type II belongs to syntactic NI. The selection between these two types of Northern Paiwan NI hinges on the constituency of the incorporated nouns, following the restriction on the incorporated spatial N\(^0\), which is realized as a morphologically-bound form, as well as the natural case-marking principle in which the incorporated spatial noun phrase/determiner phrase (NP/DP) should be obligatorily case-marked.

The paper is organized as follows. Section 2 will introduce the typological characteristics of NI and its theoretical background. Section 3 will demonstrate the typological properties of NI in Northern Paiwan. Section 4 will discuss the classification of Northern Paiwan NI and the selection between the two types of NI in Northern Paiwan. Finally, Section 5 will present the conclusion.

---

\(^3\) As opposed to the N-V word order under head movement analysis (where N results in left adjunction). Type I NI exhibits the V-N word order because there is no similar N-V constraint on word order under a lexical operation.

\(^4\) Abbreviations used in this paper are listed as follows: ABS, absolutive; ACC, accusative; AGR, agreement; BA, beneficiary applicative; COS, change of state; ERG, ergative; GEN, genitive; INTR, intransitive; IRR, irrealis; LNK, linker; NEG, negative; OBL, oblique; PRF, perfective; PL, plural; RED, reduplication; REL, relativizer; S, singular; SJTV, subjunctive; TR, transitive
2 NI typology and theoretical background

2.1 Bare noun incorporation
Incorporated nouns surface as bare nouns in most NI languages; that is, INs are realized as noun roots or stems and are devoid of the extra markings of case, determiners, and numerals (Mardirussian 1975, Rosen 1989). Mithun (1984) divided NI into four groups (Type 1 to Type 4) with respect to functional criteria. Rosen (1989) generalized these four types of NI into two word-formation processes that occur pre-syntactically in the lexicon: compounding NI, which corresponds to Mithun’s (1984) Types 1 to 3,\(^5\) as shown in (2), and classifying NI, which corresponds to Mithun’s (1984) Type 4, as shown in (3) and (4):

(2) \textit{Ni-naca-qua.}
\begin{itemize}
  \item I-flesh-eat
\end{itemize}
‘I eat flesh.’ (Nahuatl, Sapir 1911)

(3) \textit{Yedi bi-musa-tuwi-ban.}
\begin{itemize}
  \item those 1S-cat-buy-PAST
\end{itemize}
‘I bought those cat.’ (Southern Tiwa, Allen et al. 1984)

(4) \textit{...Bene-dulg-nang mangaralalymayn.}
\begin{itemize}
  \item they.two.ABS-tree-saw cashew.nut.ACC
\end{itemize}
‘…they saw a cashew tree.’ (Gunwinggu, Oates 1964)

Compounding NI usually involves a valency decrease. Typically, an incorporated noun does not receive a referential interpretation, as shown in (2). In classifying NI, the valency of the clause is not affected. A general noun is incorporated by the verb, while the clause remains transitive. In languages with classifying NI, it is possible to incorporate a noun but leave its modifiers stranded, as shown in (3), and it may also give rise to noun doubling: the independent noun in the object position is either a double of the IN or a more specific noun referring to the same element as the IN, as shown in (4). The incorporated noun bears a semantic/grammatical relation to the host verb. In most cases, the patient and theme (internal arguments) are more likely to incorporate than the external and peripheral arguments.

2.2 Noun phrase incorporation: pseudo noun incorporation
In Niuean, an Oceanic language, NI exhibits both syntactic and morphological aspects. The main verbs not only incorporate bare nouns but also complex noun phrases, as shown in (5a-b).

(5a) Ne inu kofe kono a Mele.
\begin{itemize}
  \item Past drink coffee bitter ABS Mary
\end{itemize}
‘Mary drank bitter coffee.’

(5b) According to Mithun (1984), Type 1 NI is typically lexical compounding. Type 2 NI is used to manipulate the case marking of various participants in a sentence. After NI, the direct object slot is left open, and the oblique argument \textit{ekek} can be promoted to the direct object position (ABS), while the demoted direct object \textit{kaynge} is present as an IN, as shown in (i).

(i) \textit{Tumg-e ekek kaynge-nme-nen.}
\begin{itemize}
  \item friend-ERG son (ABS) bear-kill-AGR.3S:3S
\end{itemize}
‘The friend kills a bear for the son.’ (Chukchi, Spencer 1995)

In Type 3, the IN can receive a referential interpretation, as shown in (ii).

(ii) \textit{Ti-shut-pe-ban.}
\begin{itemize}
  \item 1S-shirt-make-PAST
\end{itemize}
‘I made the/a shirt.’ (Southern Tiwa, Allen et al. 1984)
Massam (2001) proposed that as opposed to head incorporation in other languages, pseudo noun incorporation (PNI) in Niuean involves the incorporation of phrasal units, such as adjectives and certain types of relative clauses, as shown in (5b). Massam (2001) divided PNI into three types. In Type I General PNI, incorporated nouns are nonspecific and nonreferential and occur productively with an open class of verbs. Type II Existential PNI occurs with a closed class of existential verbs, and it allows the incorporated nominal to be modified by a regular relative clause. Type III PNI involves the incorporation of instrumental arguments (i.e., the English ‘go by car’ is interpreted as ‘go car’). PNI is characterized by the following features in common: the adjacency of a verb to the incorporated NP; the absence of function words indicating case and plurality for the incorporated NP; and the intransitivity of the [V+NP] complex (Massam 2001).

2.3 Interim summary

Based on the discussion above, three types of noun incorporation have been identified: compounding NI, classifying NI, and PNI. In compounding NI, the valency of the clause decreases, but in classifying NI the valency of the clause is not affected. In PNI, noun phrases can incorporate. All three types show a structural similarity—the incorporated nominals must be stripped of their usual case markings. The typological characteristics of each NI type are summarized in Table 1:

<table>
<thead>
<tr>
<th>Properties</th>
<th>Decrease in Clause Valency</th>
<th>“Patient/theme” Incorporation</th>
<th>Case Stripping</th>
<th>IN as (Modified) Noun Phrase</th>
<th>Modifier Stranding</th>
<th>Noun Doubling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compounding NI</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Classifying NI</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Pseudo NI</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

2.4 Lexical approach vs. syntactic approach

The lexical approach to NI is attributed to Sapir (1911), where NI is defined as a compositional or derivational process of compounding a noun stem with a verb. The lexical view has been accepted by Mithun (1984), Di Sciullo and William (1987), Rosen (1989), and Mohanan (1995). Lexicalists have claimed that NI is purely a lexical phenomenon that should be kept separate from syntactic processes. In this case, the NI word formation rules by which new words are produced are assigned a category label in the lexicon and are used as simple words in the syntax. Mithun (1984) pointed out that NI processes may involve lexicalization, which results in phonological and semantic idiosyncrasies that are not found in nonlexicalized syntactic constructions. Rosen (1989) proposed a unified lexicalist approach and indicated that all classifier-incorporating languages are pro-drop languages and they all freely allow pronoun-dropping in all positions. Along this line of thought, in classifying NI, null objects pro and modifiers occur independently of NI.

---

6 The relative clause is not itself incorporated but appears on the right, outside of the V-NP. See Massam (2001, 2009).
Figure 1: Pro-drop in classifying NI.

\[\text{VP} [\text{V-N}] \text{ modifier } \text{pro\_i}]\]

As shown in Figure 1, the stranding modifier together with the null pronoun occupies an argument position; that is, the V-N complex is base-generated in terms of the word formation process and thus has nothing to do with head movement (cf. Baker 1988).

The syntactic approach to NI was first proposed in the work of Sadock (1980, 1985) and Baker (1988, 1996), both of whom argued that incorporation shows the interdependency rather than the separation between syntax and morphology. Sadock (1980, 1985) claimed that the morphological component and the syntactic component can parse the same string of morphemes in different ways. Therefore, morphology can group a noun (stem) with the verb at the same time as syntax groups it with the V in a verb phrase (VP); the N and NP are related as parallel representations. Baker (1988) preferred the syntactic explanation over the morphological explanation, pointing out that the productivity and the referential transparency of NI suggest that it is a syntactic process rather than a lexical one. NI does not behave on a par with compounding, and NI differs from English compounding in two aspects: categorical status and referential transparency. The N-V compounds in English are necessarily de-verbal, whereas in NI languages (e.g. Onondaga) the N-V combination is regularly the main verb of its clause. Moreover, incorporated nouns in some languages are fully referential in a way that compounded nouns in English are not. In the syntactic approach, NI is the result of a head noun moving to the verb (or preposition) that selects or governs it. As shown in Figure 2, N (stem) and NP (argument) are related in a single syntactic operation via the movement of the head noun into the verb, leaving a trace.

Figure 2: Head movement in NI.

\[\text{VP} [\text{vN-V}] [\text{modifier} \text{ti}]\]

Baker (1988) claimed that noun incorporation is subject to the Empty Category Principle (ECP). The ECP stipulates that verbs may only incorporate those nouns whose traces are properly governed (antecedent-governed). The ECP rules out many types of NI. The most salient types of incorporation that are prohibited by the ECP are the incorporation of subjects and elements out of adjuncts. The subject is an external argument occupying the specifier (SPEC) position, so decreasing the subject to a V may impede the IN in c-commanding its trace. The same is true for adjunct incorporation, which is prohibited because an adjunct phrase does not generate a c-commanding domain and thus will block the governing of the trace by the root. While the ECP works out quite well for most cases of NI, it fails to account for some peculiar cases. In some languages, NI incorporates noun phrases as well as nouns (this issue will be revisited shortly). In Turkish and Athapaskan, NI can involve the incorporation of an agent, in violation of the ECP (Cook and Wilhelm 1998). Moreover, Chukchi incorporation does not respect Baker’s (1988, 1996) syntactic principles, as Chukchi allows the free incorporation of adjuncts, which, in Baker’s model, would lead to a violation of the ECP (Spencer 1995).

To deal with noun phrase incorporation, there are two syntactic alternatives to Baker’s (1988) head-movement analysis. Massam (2001) argued that what has been called NI in that language is simply the result of forming a verb phrase through ordinary syntactic merging. More specifically, pseudo noun incorporation is the result of the condition in which the direct object that is the first thing to merge with a V does not scramble or undergo an object shift to a position outside the minimal VP. The object remains adjacent to the verb in a very tight syntactic phrase, moving with it to the SPEC position, the tense phrase (TP), as shown in Figure 3.

Figure 3: Merge analysis in NI.

\[[\text{VP}, [\text{NP}]] \text{INFL}, \text{VP}]\]
The second syntactic alternative is to reanalyze noun incorporation as a “remnant NP movement” along the lines of Koopman and Szabolcsi’s (2000) analysis of verb cluster formation in Hungarian and Dutch (cf. phrasal movement analysis in Barrie and Mathieu (2016)). Unlike head-movement analysis, this analysis moves a full phrase and targets a SPEC position, as shown in Figure 4.

**Figure 4: Remnant NP movement in NI.**

\[ [vP \text{NP}, [v' [vV] [\text{modifier} \text{NP}_0]]] \]

### 3 Typological characteristics of NI in Northern Paiwan

In Northern Paiwan, NI shows unique properties that are not quite comparable with the NI typology reported in the literature. These special characteristics will be discussed in the following sections.

#### 3.1 INs as specific/non-specific spatial nominals

Noun incorporation in Northern Paiwan does not apply symmetrically. Activity verbs like *keman* ‘eat’ and state verbs like *tjengelay* ‘love’ do not allow noun incorporation, as shown in (6b-c) and (7b-c), where the V-N complex verb does not attract the clitic pronoun *=aken*. Only the category of spatial (motion/location/path) verbs triggers NI, as in (8a-c), where the V-N verb does attract the clitic pronoun *=aken*.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6a)</td>
<td><em>K&lt;em&gt;an=aken tua ci’aw.</em></td>
</tr>
<tr>
<td></td>
<td>eat&lt;INTR&gt;=1S.ABS OBL fish</td>
</tr>
<tr>
<td></td>
<td>‘I eat some fish.’</td>
</tr>
<tr>
<td>(6b)</td>
<td><em>K&lt;em&gt;an-ci’aw=aken.</em></td>
</tr>
<tr>
<td></td>
<td>eat&lt;INTR&gt;-fish=1S.ABS</td>
</tr>
<tr>
<td></td>
<td>Intended: ‘I eat fish.’</td>
</tr>
<tr>
<td>(6c)</td>
<td><em>K&lt;em&gt;an-tua-ci’aw=aken.</em></td>
</tr>
<tr>
<td></td>
<td>eat&lt;INTR&gt;-OBL-fish=1S.ABS</td>
</tr>
<tr>
<td></td>
<td>Intended: ‘I eat fish.’</td>
</tr>
<tr>
<td></td>
<td>(activity verb)</td>
</tr>
<tr>
<td>(7a)</td>
<td><em>Tjengelay=aken tua vatu.</em></td>
</tr>
<tr>
<td></td>
<td>love.INTR=1S.ABS OBL dog</td>
</tr>
<tr>
<td></td>
<td>‘I love a dog.’</td>
</tr>
<tr>
<td>(7b)</td>
<td><em>Tjengelay-vatu=aken.</em></td>
</tr>
<tr>
<td></td>
<td>love.INTR-dog=1S.ABS</td>
</tr>
<tr>
<td></td>
<td>Intended: ‘I love dogs.’</td>
</tr>
<tr>
<td>(7c)</td>
<td><em>Tjengelay-tua-vatu=aken.</em></td>
</tr>
<tr>
<td></td>
<td>love.INTR-OBL-dog=1S.ABS</td>
</tr>
<tr>
<td></td>
<td>Intended: ‘I love dogs.’</td>
</tr>
<tr>
<td></td>
<td>(stative verb)</td>
</tr>
<tr>
<td>(8a)</td>
<td><em>S&lt;em&gt;a-gade=aken a masengseng.</em></td>
</tr>
<tr>
<td></td>
<td>go.to&lt;INTR&gt;-mountain=1S.ABS LNK work.INTR</td>
</tr>
<tr>
<td></td>
<td>‘I (usually) go to mountain areas to work.’</td>
</tr>
</tbody>
</table>
Unlike normal cases where INs bear the semantic role of patient/theme to the host verb, in Northern Paiwan INs are restricted to peripheral arguments that denote “space” (location/goal/source), including place names, locative pronouns, directional nouns, proper names, and objective nouns (cf. Sapir 1911, Seiter 1980, Mithun 1984, Baker 1988, Johns 2009). In most NI languages, INs are generally generic and nonspecific in reference and the noun-verb complex is used to express characteristic or institutionalized activities or states. However, in Northern Paiwan, both nonspecific and specific nouns can incorporate. In (8a), the IN gade is a generic or nonspecific location, while palang in (8b) is a proper name denoting a specific personal place and gaku in (8c) is a specific location modified by a genitive phrase.

3.2 Case phrase incorporation

In addition to bare locative nouns, in Northern Paiwan, spatial verbs can incorporate case-marked personal names and objective nouns, both of which denote locations/goals/sources, as shown in (8b-c). Case phrase incorporation can be clearly distinguished from processes of verbal prefixation, where case markers are not involved, as shown in (9a-b).

Typologically, the incorporation of case phrases is quite unusual. Is the incorporated case tua/tjay really an oblique case marker in Northern Paiwan? Does the oblique case marker really incorporate? The answers are positive for the following reasons: (i) they are inflected on a par with regular case markers; (ii) their distribution is parallel to that of regular case markers; and (iii) “wh-extraction” is not attested in NI.

As shown in (10a-b), typically, the oblique case markers tua/tjay indicate the relation between verbs and arguments in intransitive constructions (cf. Tang et al. 1998), and they may mark the theme argument vavu’a selected by the verb veneli, as shown in (10a), or the patient kaljalju selected by penangulj, as shown in (10b). The same is true for spatial verbs in Northern Paiwan.
In the NI structure, the intransitive motion verb *sema* selects an oblique location (goal) argument, as shown in (11a), in the same way as the INTR active verb *veneli* selects an oblique locative theme, as shown in (10a). Moreover, the form of the oblique case marker will shift according to nominal properties—whether they are proper nouns or not. *Tua* co-occurs with common nouns, as shown in (10a), while *tjay* co-occurs with proper nouns, as shown in (10b). The sensitivity to nominal properties can be observed in the arguments of both normal verbs and spatial verbs, as well as the independent forms, as shown in (10a-b), and the incorporated forms, as shown in (11a-b).

Unlike normal oblique arguments selected by lexical verbs, the case-marked incorporated noun cannot undergo cleft-formation to form wh-sentences, as shown by the following contrasts in (12a-c) and (13a-c).

(12a)  *Na-p<en>*-angulj=aken  tjay palang.
PRF-beat<INTR>=1S.ABS  OBL  Palang  
‘I beat Palang.’

(12b)  *Na-p<en>*-angulj=esun tjay ima?
PRF-beat<INTR>=2S.ABS  OBL  who  
‘Who did you beat?’  (wh-in situ)

(12c)  *Timai  a  [DP [CP opi [TP [vP su-p<en>*angulj ti ]]]]? 
who  ABS  2S.ERG-beat<PRF.TR>  
‘Who did you beat?’  (wh-cleft)

(13a)  *Na-k<em>*-si-tjay-palang=aken.
PRF-be.from<INTR>-OBL-Palang=1S.ABS  
‘I was from Palang’s (place).’

(13b)  *Na-k<em>*-asi-tjay-ima=esun.
PRF-be.from<INTR>-OBL-who=2S.ABS  
Lit. ‘Who did you come from?’/ ‘Who’s place did you come from?’  (wh-in situ)

(13c)  *Timai  a  su-k<en>*-asi 
who  ABS  2S.ERG-be.from<PRF.TR>  
(wh-cleft)

If a normal oblique argument is in question, it can either stay in situ, as shown in (12b), or undergo cleft-formation through “extraction”, as shown in (12c) (Tsai 1997, 2003). However, the “wh-extraction” does not apply to the oblique incorporated noun. In (13a), the oblique argument *palang* cannot be wh-extracted and the wh-element must stay in situ. The host verb and the incorporated noun together with its case marker constitute a large and unbreakable V0 unit, and it thus can attract pronominal clitics, as shown in (13a-b).

The phenomena of case phrase (KP) incorporation poses challenges for both traditional lexical and syntactic approaches. Lexical analysis is not adequate for KP incorporation because case markers should not occur in V-N compounds. Case phrase incorporation violates head-movement analysis, which is applicable to most NI because it incorporates elements larger than N0 and it does not deprive it of the case.
3.3 “Adjunct” incorporation

It is noteworthy that KP incorporation in Northern Paiwan involves “adjunct” incorporation, which violates the ECP that head movement should observe. This argument comes from the observation that Northern Paiwan is a language that has an ergative nature. In Northern Paiwan, intransitive verbs are identified as antipassive. According to Chang (2004), the grammatical object is systematically absent in the case system of Northern Paiwan. Various semantic/syntactic tests have shown that the “object” (tua/tjay-marked OBL argument) is in fact denoted, as would be expected under an ergative analysis. First, the oblique case markers tua/tjay not only encode patient/theme arguments, they also mark peripheral arguments. In (14a), tjay is marked for the beneficiary and in (14b) for the cause.

(14a) \[ M\text{-}alap=aken tua zua a kava tjay palang. \]
\[
\text{INTR-take=1S.ABS OBL that LNK clothes 3S.OBL Palang}
\]
‘I got the clothes for Palang.’

(14b) \[ Ma\text{-}zeli ti Palang tjay kaljalju. \]
\[
\text{INTR-tired abs Palang OBL Kaljalju}
\]
‘Palang is tired because of Kaljalju.’

(14c) \[ K<\text{em}>=an=aken tua ci’aw. \]
\[
\text{eat<INTR>=1S.ABS OBL fish}
\]
‘I ate some fish.’

Second, the OBL argument is by default nonspecific or indefinite, as shown in (14c), as has been reported in many ergative Austronesian languages (such as Tagalog and Kavalan) (Liao 2002, Paul and Travis 2006). Third, Northern Paiwan lacks object-control constructions (Chang & Tsai 2001). In English, an object control construction targets a direct object. In other words, the OBL argument can control the pro in the embedded clause, as shown in (15a). In Northern Paiwan, however, what has been taken as an object-control construal in English turns out to be a genitive/subject-control construction. In (15b), the obligatory occurrence of the causative pa- in the embedded clauses indicates that the agent aken rather than the object palang serves as the controller.

(15a) John persuaded Maryi [Proi to swim.]

(15b) \[ Pa’alid=aken, tjay palang, \]
\[
\text{force.INTR=1S.ABS OBL Palang LNK CAUS-eat<INTR>}
\]
\[\text{kan/*k<\text{em}}>=an tua kinsa]. \]
\[
\text{CAUS-eat<INTR> OBL cooked.rice}
\]
‘I forced Palang to eat cooked rice.’

---

7 One of the anonymous reviewers proposed that “adjunct” incorporation here may actually involve prepositional phrase (PP) incorporation with an oblique (OBL) marker. In the reviewer’s view, a PP is selected by a verb of location or movement and occupies the direct object position of location or movement verbs. In Niuean, for example, there is (mandatory) incorporation of PPs with location verbs. Massam (2009) considered PPs as sisters to Vs, hence, moving along with the verb to the initial position. I have certain reservations about this analysis because whether there is a distinct PP category in Northern Paiwan is still an unsolved issue. Most of the previous studies have shown that Paiwan, as well as other Formosan languages, lack the PP category (Starosta 1988, Wu 2004). To solve this problem, more investigations and reassessments on the existence of the PP category in Northern Paiwan are required.
Another piece of evidence in support of ergative analysis is that in transitive constructions the agent (A) argument has a certain syntactic prominence and can serve as the binder of its reflexive pronoun in the subject position (Aldridge 2004, 2008). As shown in (15c), the ergative DP palang binds the absolutive anaphor macizi. This particular binding phenomenon that is not attested in accusative languages (like English) but takes place in ergative languages exemplifies the ergative nature of Northern Paiwan.

4 Lexical NI vs. syntactic NI

In Northern Paiwan, two types of noun incorporation can be distinguished in terms of different selections of incorporated nominal elements and differences in lexical and morphosyntactic behaviors. In this section, I will argue that these two types of Northern Paiwan NI are associated with separate grammatical processes—lexical NI is morphologically derived and syntactic NI is syntactically derived.

4.1 Lexical NI

4.1.1 IN as a N\textsuperscript{0} nominal root

In the lexical type, IN includes place names, such as tjimur in (16a), location nouns, such as gade in (16b), directional nouns, such as vavaw in (16c), and locative pronouns, such as maza in (16d). These nouns are morphologically realized as bound nominal roots in Northern Paiwan (Chang 2007).

(16a) \textit{Uru=s<em>a-tjimur=aken.}\n\textit{IRR=go.to<INTR>-Tjimur=1S.ABS}\n\textit{‘I will go to Tjimur.’}

(16b) \textit{Pasa-gade=amen a vaik nukatjatiaw.}\n\textit{move.to.INTR-mountain-1PL.ABS LNK go.INTR everyday}\n\textit{‘We go to mountain areas every day.’}

(16c) \textit{S<em>a-vavaw=aken a m-alap tua kava.}\n\textit{go.to<INTR>-top=1S.ABS LNK INTR-take OBL clothes}\n\textit{‘I went up to take the clothes.’}

(16d) \textit{K<em>asi-maza=aken a tjalu-zua.}\n\textit{be.from<INTR>-here=1s.abs LNK go.toward-there.INTR}\n\textit{‘I move from here to there.’}

It has been observed that in Northern Paiwan, these spatial nouns cannot occur as independent nominal arguments (Chang 2007). Instead, they appear as bound forms that cannot co-occur with case markers, as shown in (17a), and cannot serve as nominal modifiers, as shown in (17b-d).

(17a) \textit{Tjengeljay=aken tua *(i)-tjimur/*(i)-gade.}\n\textit{like.INTR=1S.ABS OBL be.in-Tjimur/be.in-mountain}\n\textit{‘I like Tjimur/mountains.’}
(17b) Maca’u   a   *(kasi)-tjimu a kakeDian.
smart.INTR ABS be.from-Tjimur REL children
‘The children from Santimen are smart.’

(17c) Nguangua’ a *(i)-vavaw a uma’.
beautiful.INTR ABS be.in-top REL house
‘The house on the top is beautiful.’

(17d) Ini=aken a tjengeljay tua *(i)-gade a sinsi.
NEG=1S.ABS LNK like (INTR) OBL be.in-mountain REL teacher
‘I don’t like teachers in the mountain.’

In contrast, the normal objective/personal nominals like vatu and kaljalju must co-occur with case markers when appearing as independent arguments, as shown in (18a-b), and can serve as nominal modifiers, as shown in (18c-d).

(18a) P<en>angulj=aken *(tua) vatu.
beat<INTR>=1S.ABS OBL dog
‘I beat a dog.’

(18b) P<en>angulj =aken *(tjay) kaljalju.
beat<INTR>=1S.ABS OBL Kaljalju
‘I beat Kaljalju.’

(18c) T<em>ekel=aken tua kuka a siaw.
drink<INTR>=1S.ABS OBL chicken REL soup
‘I drank the chicken soup.’

(18d) Tjengeljay=aken tua kava ni camak.
like.INTR=1S.ABS OBL clothes GEN Camak
‘I like Camak’s clothes.’

Systematically, spatial nouns in Northern Paiwan tend to surface as N°, which cannot be case-marked, and respect the N° restriction (Chang 2007) in (19).

(19) N° restriction:
N° nominals in Northern Paiwan cannot be case-marked.
The \( N^0 \) nominals therefore morphologically are incorporated into the spatial verbs that select them, as shown in Figure 5.

**Figure 5: Selection of lexical NI**

\[
V^0 + N^0 \\
\| \quad \| \\
\text{spatial verb} \quad \text{spatial nominal root}
\]

**Semantic restriction:** selectional restriction

**Grammatical restriction:** combination of \( V^0 \) and \( N^0 \)

### 4.1.2 Lexical NI and Idiosyncrasies

Lexical NI may exhibit lexical idiosyncrasies. A lexical NI verb can be further lexicalized, and the meaning becomes nontransparent and cannot be predicted by its appearance. Consider the following examples in (20a-b) and (21a-b).

(20a)  
\[i-zua=aken\]  
\[a \quad l<em>anguy.\]  
\[\text{be.in.INTR-there=1S.ABS LNK swim<INTR>}\]  
\[\text{‘I swam there.’}\]

(20b)  
\[Izua\]  
\[a \quad ku-paisu.\]  
\[\text{exist.INTR ABS 1S.GEN-money}\]  
\[\text{Lit. ‘My money exists.’ / ‘I have money.’}\]

(21a)  
\[i-vavaw=aken\]  
\[\text{be.on-top.INTR=1S.ABS}\]  
\[\text{‘I am on the top.’}\]

(21b)  
\[Tja-iwaw=aken\]  
\[\text{more-tall.INTR=1S.ABS 2S.OBL}\]  
\[\text{‘I am taller than you.’}\]

In (20a) the lexical NI verb \( i-zua \) denotes the existence of a location, and it can be further lexicalized as a common existential verb meaning ‘to have’ in English, as shown in (20b). In (21a-b), the lexical NI verb \( i-vavaw \) ‘be on the top’ can be lexicalized as the static verb \( ivawav \) ‘be tall’. The lexical idiosyncrasies attested in Northern Paiwan provide another piece of evidence, showing that lexical NI can be morphologically derived.

### 4.1.3 V-N order

In Northern Paiwan, a lexical NI verb always surfaces as a V-N complex verb. The V-N word order is not predicted via head movement. As shown in Figure 2, head-movement analysis should predict that a lexical NI verb complex occurs in an N-V order, where the N results in left adjunction. The V-N word order shows that lexical NI can be morphologically derived because there is no similar N-V constraint on word order under a lexical operation.
4.1.4 Full lexical status

The lexical NI construction can be simply identified as a [V-N] compounding verb. Like a normal lexical verb, a lexical NI verb can undergo a full range of morphosyntactic processes. For one thing, a lexical NI verb can be inflected for voice, and it serves as an intransitive or transitive verb in the same way as a lexical verb does. There are two intransitive subtypes that can be observed. In the normal intransitive type, a nonspecific spatial noun is incorporated, and the NI gives rise to a habitual reading, as shown in (22a). In the special intransitive type, the lexical NI containing the locative pronouns *zua* ‘there’ and *maza* ‘here’ can be doubled by a free-standing place noun, as shown in (22b), a phenomenon attested in classifying NI. In (22b), the valency of the clause is not affected. The motion/location verbs incorporate the locative pronouns *zua/maza*, which denote generic notions of location, and they are coreferential with the definite locative noun *pana* in prepositional phrases.

(22a) \( \text{Pasa-}gaku=\text{amen} \quad a \quad \text{vaik.} \)
\[
\begin{array}{ll}
\text{move.to.INTR-school-1PL.ABS} & \text{LNK} \\
\text{go.INTR} & \\
\end{array}
\]
‘We go to school (for study every day).’

(22b) \( \text{Uru-}s<\text{em}>a-\text{zua/maza}=\text{aken} \quad i \quad \text{tua} \quad [\text{BC nguangua'} \quad a] \quad \text{pana.} \)
\[
\begin{array}{llll}
\text{IRR-go.to<INTR>there/here=1S.ABS} & \text{be.at} & \text{OBL} & \text{beautiful} \\
\text{REL} & \text{river} & & \\
\end{array}
\]
‘I will go to the beautiful riverbank (there/here).’

A lexical NI verb taking the transitive marker -en and the applicative marker *si-* may introduce a new absolutive theme and beneficiary arguments, as shown in (23a) and (23b), respectively.

(23a) \( \text{Uru=}k\text{u-pasa-pana-en} \quad a \quad \text{’acilay}. \)
\[
\begin{array}{ll}
\text{IRR=1S.ERG-move.to-river.bank-TR} & \text{ABS} \\
\text{stone} & \\
\end{array}
\]
‘I will move the stone to river bank.’

(23b) \( \text{Uru=}k\text{u-si-pasa-pana} \quad \text{tua} \quad \text{’acilay ti kaljalju}. \)
\[
\begin{array}{llll}
\text{IRR=1S.ERG-BA-move.to-river.bank} & \text{obl} & \text{stone} & \text{ABS} & \text{Kaljalju} \\
\end{array}
\]
‘I will move stones to river bank for Kaljalju.’

This transitive/intransitive voice alternation is widely attested in Northern Paiwan lexical verbs, as shown in (24a-c), where the lexical verb ‘eat’ can be inflected for the intransitive/transitive/applicative:

(24a) \( \text{Uru=}k<\text{em}>a\text{=}\text{aken tua kinsa}. \)
\[
\begin{array}{ll}
\text{IRR=eat<INTR>=1S.ABS} & \text{OBL} \\
\text{cooked.rice} & \\
\end{array}
\]
‘I will eat cooked rice.’

(24b) \( \text{Ku-}k\text{an-en a kinsa}. \)
\[
\begin{array}{ll}
\text{1S.ERG-eat-TR} & \text{ABS} \\
\text{cooked.rice} & \\
\end{array}
\]
‘I will eat the cooked rice.’

(24c) \( \text{Ku-si-}k\text{an tua kinsa ti kaljalju}. \)
\[
\begin{array}{llll}
\text{1S.REG-BA-eat} & \text{OBL} & \text{cooked.rice} & \text{ABS} & \text{Kaljalju} \\
\end{array}
\]
‘I will eat cooked rice for Kaljalju.’

There is an additional reason for this phenomenon: a lexical NI verb can be fully-inflected for tense or aspect like a lexical verb, and it can co-occur with the irrealis marker, the perfective marker, and the change-of-state marker, as well as undergo the process of reduplication indicating the progressive, as shown in the comparison between the lexical NI verbs in (25a-d) and the lexical verbs in (26a-d):
(25a)  \textit{Uru}=s\textless em\textgreater a-agde=aken.  
\hspace{1cm} \text{IRR}=go\textless INTR\textgreater -mountain=1S.ABS  
\hspace{1cm} ‘I will go to mountain areas.’  \hspace{1cm} \text{Irrealis}

(25b)  \textit{Na-k}\textless em\textgreater asi-pairang=aken.  
\hspace{1cm} \text{prf-be.from}\textless INTR\textgreater -plain=1S.ABS  
\hspace{1cm} ‘I came from the plain.’  \hspace{1cm} \text{Perfective}

(25c)  \textit{Tjalu-pairang=anga=amen.}  
\hspace{1cm} \text{arrive-plain=COS=1PL.ABS}  
\hspace{1cm} ‘We have already arrived at the plain.’  \hspace{1cm} \text{Change-of-state}

(25d)  \textit{S}\textless em\textgreater a-pana-pana=aken.  
\hspace{1cm} \text{go.to }\textless \text{INTR}\textgreater -river-RED=1S.ABS  
\hspace{1cm} ‘We have already arrived at the plain.’  \hspace{1cm} \text{Progressive}

(26a)  \textit{Uru}=k\textless em\textgreater an=aken tua \textit{ci’aw}.  
\hspace{1cm} \text{IRR}=eat\textless \text{INTR}\textgreater =1S.ABS OBL fish  
\hspace{1cm} ‘I will eat fish.’  \hspace{1cm} \text{Irrealis}

(26b)  \textit{Na-k}\textless em\textgreater an=aken tua \textit{ci’aw}.  
\hspace{1cm} \text{prf-eat}\textless \text{INTR}\textgreater =1S.ABS OBL fish  
\hspace{1cm} ‘I ate fish.’  \hspace{1cm} \text{Perfective}

(26c)  \textit{K}\textless em\textgreater an=anga=aken tua \textit{ci’aw}.  
\hspace{1cm} \text{eat}\textless \text{INTR}\textgreater =COS=1S.ABS OBL fish  
\hspace{1cm} ‘I have already eaten fish.’  \hspace{1cm} \text{Change-of-state}

(26d)  \textit{K}\textless em\textgreater a-kan=aken tua \textit{ci’aw}.  
\hspace{1cm} \text{RED-eat}\textless \text{INTR}\textgreater =1S.ABS OBL fish  
\hspace{1cm} ‘I am eating fish now.’  \hspace{1cm} \text{Progressive}

To sum up, Lexical NI is morphologically derived and therefore it bears strong lexical properties. First, the incorporated spatial noun is a N0 nominal root, combined with its V0 host verb. Second, idiosyncrasies can be observed in the NI process. Moreover, the lexical NI verb exhibits the V-N word order, which is not predicted under head-movement analysis. Finally, the lexical NI verb can be fully inflected like a lexical verb.

\section*{4.2 Syntactic NI}

\subsection*{4.2.1 IN as a case-marked NP/DP nominal}
Syntactic NI involves phrasal incorporation. In this type, a host verb can incorporate objective nouns or personal names that indicate location/goal/source, together with the oblique case markers \textit{tja} and \textit{tua}, as shown in (27a-b). The constituent of the IN can be further extended. It is possible for an IN to contain other modifiers like relative clauses and genitive phrases as well, as shown in (27c-d).
Why do the oblique case markers tua and tjay incorporate? In contrast to lexical NI, the INs in syntactic NI have a functional projection higher than N0—they are noun phrases or determiner phrases, and they must always co-occur with case markers when appearing in the argument position. When appearing with normal lexical verbs, a personal/objective nominal (NP), as shown in (28a-b), and a modified nominal (DP), as shown in (28c-d), are obligatorily marked by the oblique case tjay/tua in the intransitive form:

(28a) P<en>angulj=aken *(tjay) [NP palang].
beat<INTR>=1S.ABS OBL Palang
‘I beat Palang.’

(28b) V<en>eli=aken *(tua) [NP uma’].
beat<INTR>=1S.ABS OBL house
‘I bought a house.’

(28c) Tjengeljay=aken *(tua) [DP uma’- [RC a v<in>eli ni camak]].
like.INTR=1S.ABS OBL house REL buy<PRF> GEN Camak
‘I like the house that Camak bought.’

(28d) Tjengeljay=aken *(tua) [DP uma’ [ni camak]].
like.INTR=1S.ABS OBL house GEN Camak
‘I like Camak’s house.’

In other words, in Northern Paiwan, NP/DP must follow the case-marking restriction shown in (29), and this restriction also applies to the incorporated nominals in syntactic NI (Chang 2007):

(29) Case-marking Restriction:
Every NP/DP must be case-marked when occurring in the argument position in Northern Paiwan.

As a result, the oblique case-marked NP/DP-incorporated nominals are syntactically selected by the host spatial verb. The selection of syntactic NI can be formulated as in Figure 6.
Figure 6: Selection of syntactic NI.

\[
V^0 + K + \text{NP/DP} \\
\begin{array}{c|c|c}
\text{Spatial verb} & \text{tua/tjay} & \text{objective noun; personal name/objective noun; personal name being modified}
\end{array}
\]

**Thematic restriction**: IN must be objective and personal arguments representing space (including goal, source, and location)

**Case-marking Restriction**: Every NP/DP must be case-marked

The selection of a syntactic NI follows the thematic restriction imposed on the NP/DP by spatial verbs and the case-marking restriction on the NP/DP. A spatial verb may c-select an NP, an objective noun, or a personal name, or a DP, an objective noun, or a personal name containing its modifiers, such as relative clauses and possessive phrases. In Northern Paiwan, every NP/DP must be obligatorily case-marked when they occur in argument positions. As aforementioned, *tua* is used for (modified) common objective nouns and *tjay* for personal proper names.

4.2.2 Modifier stranding
As aforementioned, syntactic NI shows that “more than” a noun can incorporate. Modifiers like genitive phrases and relative clauses can appear with INs, as shown in (30a-b):

(30a) \[ Uru-s<em>a-tua-[uma’ \ [a v-in-eli ni camak]]=aken. \]
IRR-go.to<INTR>-OBL-house REL buy-TR GEN camak =1S.ABS
‘I will go to the house that Camak bought.’

(30b) \[ Uru-s<em>-a-tua- [uma’ [ni camak]]=amen. \]
IRR-go.to<INTR>-OBL-house GEN camak =2PL.ABS
‘We will go to Camak’s house.’

It can be seen that the clitic =aken/=amen in (30a-b) may move left of the case-marked noun and leave stranded modifying elements, as shown in (31a-b). This involves “modifier stranding” attested in classifying NI.

(31a) \[ Uru-s<em>-a-tua-uma’=aken \]
IRR-go.to<INTR>-OBL-house=1S.ABS
‘I will go to the house that Camak bought.’

(31b) \[ Uru-s<em>-a-tua-uma’=amen \]
IRR-go.to<INTR>-OBL-house=2PL.ABS
‘We will go to Camak’s house.’

---

8 This is somehow similar to pseudo noun incorporation found in Niuean (Massam 2001) in that the incorporated nominals are noun phrases rather than bare nouns. However, there is a salient difference. Unlike PNI, INs must always be associated with the oblique case markers *tua* and *tjay*.  

112
The incorporated nominal always targets the nominal head, leaving the modifiers stranded in the argument position, but not vice versa. For example, in (31b), the host verb incorporates the nominal head *uma* ‘house’, stranding the possessor *ni camak*. However, the incorporation of the modifier *camak* is prohibited, as shown in (31c).

According to Rosen (1989), the example in (31b) could be identified as a classifier NI and the syntactic NI verb *sema-tua-uma* can be lexically-derived: the stranded modifier *ni camak* must go with the null pro, which is coreferential with the NI *uma*, as shown in (32):

(32)  
\[
\text{Uru-s<em>a-tua- uma'}=\text{amen pro}_{i} \text{ ni camak.}  \\
\text{IRR-go.to\textless INTR\textgreater -OBL-house=2PL.ABS GEN Camak}  \\
\text{‘We will go to Camak’s house.’}
\]

However, lexical analysis is not tenable because with the incorporation of the oblique marker *tua*, the incorporated nominal is a phrasal unit rather than a bare noun. Moreover, there is no solid evidence to show that Northern Paiwan is a pro-drop language and the similar pro-drop construction is not attested. If the lexical analysis is on the right track, the pro in (32) should represent an independent argument outside the NI verb complex. However, the pro and its antecedent are generated within the same DP, as shown in (30b), and so the pro in (32) cannot be an independent argument. Moreover, in (33) the null external pro is unrecoverable:

(33)  
\[
\text{*Uru-s<em>a-tua- uma'}=\text{amen tua uma'} \text{ ni camak.}  \\
\text{IRR-go.to\textless INTR\textgreater -OBL-house=2PL.ABS OBL house GEN Camak}  \\
\text{Intended: ‘We will go to Camak’s house.’}
\]

These facts indicate that the stranded modifier *ni camak* might not accompany a null independent argument, and thus the NI structure in (32) cannot be lexically derived.

### 4.2.3 Semi-lexical status

Unlike lexical verbs and lexical NI verbs, the syntactic NI verb complex [V-K-NP/DP] cannot fully undergo morphosyntactic processes specified for a lexical verb. This fact reveals that the syntactically-derived verb complexes are not fully associated with lexical properties. A [V-K-NP] verb complex can be transitivized and applicativized, as shown in (34a-b); however, a [V-K-DP] verb complex, where the INs appear with modifiers, cannot be transitivized and applicativized, as shown in (35a-b):

(34a)  
\[
\text{Ku-pasa-tjay-palang-\text{en} a kava.}  \\
\text{1S.ERG-move.to-obl-Palang-TR ABS clothes}  \\
\text{‘I took the clothes to Palang’s (place).’}
\]

(34b)  
\[
\text{Ku-si-pasa-tjay-palang tua kava ti kaljalju.}  \\
\text{1S.ERG-BA-move.to-OBL-Palang OBL clothes ABS Kaljalju}  \\
\text{‘I took the clothes to Palang’s (place) for Kaljalju.’}
\]

(35a)  
\[
\text{*Ku-pasa-tua- [nguangua’ a- uma’-\text{en} a ku-kava.}  \\
\text{1S.ERG-move.to-OBL-beautifulREL-house-TR ABS 1S.GEN-clothes}  \\
\text{Intended: ‘I took my clothes to the beautiful house.’}
\]
(35b) *Ku-si-pasa-tua- [nguangua’ a]- uma’ tua ku-kava Timadju.
1S.ERG-BA-move.to-OBL-beautifulREL-house OBL 1S GEN-clothes 3S.ABS
Intended: ‘I took my clothes to the beautiful house for him.’

In addition, a syntactic NI verb complex cannot be fully inflected for tense/aspect. Although the [V-K-NP/DP] verb complex can appear with the irrealis marker *uru= and the perfective marker *na- as shown in (36a-b), it cannot co-occur with the change-of-state marker =anga, as shown in (36c-d), and cannot undergo the process of reduplication, as shown in (37a-b).

(36a) Uru=s<em>a-tjay-palang timadju.
IRR=go.to<INTR>-obl-Palang 3S.ABS
‘I will go to Palang’s (place).’ (Irrealis)

(36b) Na-s<em>a-tua-[nguangua’ a]-uma’=aken.
PRF-go.to<INTR>-OBL-beautiful REL-house=1S.ABS
‘I went to the beautiful house.’ (Perfective)

(36c) *
S<em>a-tjay-palang-anga timadju.
go.to<INTR>-OBL-Palang-COS 3S.ABS
Intended: ‘He has gone to Palang’s (place).’ (Change-of-state)

(36d) *
Pasa-zua-uma’-ni-cama-camak=aken.
mov.to.INTR-house-GEN-RED-Camak=1S.ABS
Intended: ‘I am going to Camak’s house.’ (Progressive)

(37a) *
Pasa-uma’-ni-cama-camak=aken.
mov.to.INTR-house-GEN-RED-Camak=1S.ABS
Intended: ‘I am going to Camak’s house.’ (Progressive)

To express the meanings of ungrammatical (35a) and (37b), the mechanism of “noun doubling” in lexical NI must be employed. Motion/location/path verbs must incorporate the locative pronoun *zua ‘there’ to form lexical NI verb compounds that can be inflected for voice and aspect, and the modified locative noun phrase appears again in prepositional phrases, as shown in (38a-b).

(38a) Ku-pa-sa-zua-en a ku-kava i tua [nguangua’ a] uma’
1S.GEN-CAUS-go.to-there-TR ABS 1S.GEN-clothes in OBL beautiful REL
house ‘I took my clothes to the beautiful house.’ (for (35a))

(38b) Pasa-zua-zua=aken i tua uma’ ni camak.
mov.to.INTR-there-RED=1S.ABS in OBL house GEN Camak
‘I am going to Camak’s house.’ (for (37b))
These differences in morphosyntactic behaviors can be attributed to the discrepancy in NI structures. In the [V-K-NP] verb complex, the NP is an argument without any nominal modifiers, while in [V-K-DP] the DP is an extended noun argument modified by genitive phrases or relative clauses. In general, the [V-K-NP] verb complex, which consists of three units, is more lexical-like, while the [V-K-DP] complex, which consists of more than three units, has the least lexical properties.

4.3 Summary
This study revealed that Northern Paiwan has two types of NI, which are derived separately in grammar. Generally speaking, lexical NI is morphological derived. Idiosyncrasies can be observed in lexical NI and lexical NI verbs can fully undergo morphosyntactic processes specified for a lexical verb. The INs are realized as N⁰, which is morphologically selected by the host verbs. Syntactic NI should be formed in syntax since it involves the incorporation of case markers and modifier stranding. A syntactic NI verb complex cannot be fully inflected, as the INs are realized as NPs/DPs, which are syntactically selected by the host verbs and therefore must be case-marked. The selection between lexical NI and syntactic NI hinges on the constituency of the incorporated nouns. Table 2 summarizes some crucial properties that distinguish lexical NI, syntactic NI, and other types of NI cross-linguistically.

Table 2: Cross-linguistic comparisons of NI.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Decrease in Clause Valency</th>
<th>“Patient/theme” Incorporation</th>
<th>“Adjunct” Incorporation</th>
<th>Case Stripping</th>
<th>NP/DP Incorporation</th>
<th>Modifier Stranding</th>
<th>Noun Doubling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compounding NI</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Classifying NI</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Pseudo NI</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Lexical NI</td>
<td>Yes/No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Syntactic NI</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

In general, lexical NI and syntactic NI can fit into the NI typology because in these types of NI the host verbs and the INs are all correlated via relations of s-selection. In regular NI (compounding NI, classifying NI and Pseudo NI), the IN is usually a patient/theme argument selected by a V⁰. In lexical NI and syntactic NI, the IN is a location/source/goal argument selected by a spatial V⁰. The verb complex derived from lexical NI can somehow be treated on a par with the one derived from compounding NI and classifying NI, as they have some features in common. Structurally, the derived verb complexes of the three types are truly lexical verbs. Lexical NI is similar to compounding NI in that they both involve a decrease in clause valency and case stripping but do not involve modifier stranding. Moreover, one type of lexical NI (with the incorporation of locative pronouns) is similar to classifying NI in three aspects: they both do not involve a decrease in clause valency, but case stripping and noun doubling are attested. In spite of these similarities, lexical NI differs from compounding NI and classifying NI in one major typological property: lexical NI is not associated with patient/theme incorporation. Syntactic NI and pseudo NI both involve phrasal incorporation. However, syntactic NI entertains the incorporation of case, but pseudo NI does not.
5 Conclusion
In this paper, morphosyntactic properties and the behaviors of lexical and syntactic NI in Northern Paiwan have been examined and the underlying mechanisms that determine the NI selections have been discussed. The phenomena of NI show that Northern Paiwan is a polysynthetic language since NI is a major polysynthetic property (Baker 1996). However, there are some critical issues that remain unanswered and need to be explored. How is syntactic NI derived? What syntactic operation(s) is/are involved in the process: head-to-head movement or phrasal movement (PNI merge or remnant movement)? It would be very difficult to maintain the head-movement analysis of these data since the incorporated nominals are case-marked NP/DP rather than bare N°. PNI analysis is not adequate because syntactic NI involves the incorporation of case. Phrasal movement analysis seems more convincing, but there are problems, too, with it. The major problem is that the incorporated oblique NP/DP would have to move from an adjunct position to the specifier position of the VP, in contrast to Figure 4. If the incorporated NP/DP were a direct object realized as a PP with an oblique case, then remnant-movement analysis would be quite feasible. However, more evidence should be provided to support PP analysis, since the existence of the PP category needs to be identified, and to explain why an ergative language such as Northern Paiwan has a PP goal/locative object. In addition, the derived [V-K-NP/DP] word order should be further explained. More investigations are required, leaving these questions open for further study and research.

References
Chang, Henry Yungli. 2004. AF verbs: transitive, intransitive, or both? In Studies on Sino-Tibetan languages: papers in honor of Professor Hwang-Cherng Gong on his seventieth birthday, ed. by Ying-chin Lin, Fang-min Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang, and Dah-an Ho, 95-120. Taipei: Institute of Linguistics, Academia Sinica.


MALAY VERBAL REDUPLICATION WITH THE məN- PREFIX

Meng Yang & Deborah J.M. Wong
University of California, Los Angeles
mengyang@ucla.edu & deborah.jm.wong@ucla.edu

Abstract
Malay is known for having both full reduplication (e.g. buŋa ‘flower’, buŋa-buŋa ‘flowers’) and nasal substitution when a nasal is followed by a voiceless stop. These two processes interact in verbal reduplication, where the məN- prefix attaches either to the first or second component of the reduplicated verb. When məN- attaches to the first component, the copy that doesn’t bear the prefix surfaces with a nasal homorganic to the underlying voiceless stop (e.g. mənari+nari ‘dance (continuous)’). When the məN- attaches to the second component, the copy that doesn’t bear the prefix surfaces faithfully with the underlying voiceless stop (e.g. tari+mənari ‘dance (reciprocal)’). This pattern is difficult to account for under any derivational analysis. We propose that this pattern of reduplication is best accounted for in Parallel OT. Our OT analysis includes a new markedness constraint *N_{[word-initial]}, which reflects the phonotactics of Austronesian languages of Southeast Asia.

Keywords: Malay, reduplication, məN- prefix
ISO 639-3 codes: may/msa

1 Introduction
This paper looks at the interaction between two common processes found in Malay: reduplication and nasal substitution. Reduplication is a highly productive process in Malay and enables the formation of new words from root words. According to an early study of reduplication in Malay by Asmah (1975), the process can be classified into six categories based on differences in the morphological changes of the word. The six categories are as listed in (1).

(1) a. Whole-word reduplication
   b. Rhyming and chiming reduplication
   c. Partial reduplication
   d. Reduplication with ‘em’ infixation
   e. Reduplication with root that does not stand alone
   f. Reduplication with affixation (prefix, suffix, circumfix) process

The reduplicative pattern we are interested in is (1f), reduplication with affixation, specifically the affixation of məN-. In this type of reduplication, the reduplicative process can be seen interacting with nasal substitution at the morpheme boundary between the stem and the prefix when the stem begins with a voiceless stop. When the verb is reduplicated, məN- can attach either to the first component of the reduplicated verb, giving it a continuous meaning, or to the second component of the reduplicated verb, giving it a reciprocal meaning. We show that the pattern arising from the interaction between the processes, morphological and phonological, cannot be accounted for under any derivational account. Instead, we propose that this interaction is best accounted for in a simple Optimality-Theoretic (OT) account using the assumptions of a basic parallel derivation as laid out in Prince and Smolensky (1993/2004).

1 We would like to thank Kie Zuraw and the audiences of the UCLA Phonology Seminar, the Austronesian Formal Linguistics Association (AFLA-25), and the International Symposium on Malay/Indonesian Linguistics (ISMIL-22) for their valuable feedback and suggestions. Many thanks to our consultants, Leila Louise Fitton, Sarah Amalina Ab Wahid, and Yasmin Rizal for the data and judgments. All remaining errors are our own.

Copyright vested in the author; Creative Commons Attribution License
The paper is organized as follows: In Section 2, we present the data and describe the phenomenon. In Section 3, we show the problem that the data poses for derivational accounts, as well as for McCarthy and Prince’s (1995) Parallel OT analysis. Drawing from Malay phonotactics, we then propose a revised OT analysis in Section 4. Finally, we provide a conclusion for the paper.

2 The Data

2.1 Full reduplication and məN- prefixation in Malay

In full reduplication, the entire word, whether simplex or complex, undergoes reduplication. In Malay, this process occurs with different classes of words and serves many different functions as shown in the examples in (2).2

(2) Full reduplication examples

<table>
<thead>
<tr>
<th>Word</th>
<th>Gloss</th>
<th>Reduplicated Word</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>buŋa</td>
<td>‘flower’</td>
<td>buŋa-buŋa</td>
<td>‘flowers’</td>
</tr>
<tr>
<td>oɾɑŋ</td>
<td>‘person’</td>
<td>oɾɑŋ-oɾɑŋ</td>
<td>‘scarecrow’</td>
</tr>
<tr>
<td>lambat</td>
<td>‘slow’</td>
<td>lambat-lambat</td>
<td>‘slowly’</td>
</tr>
<tr>
<td>apa</td>
<td>‘what’</td>
<td>apa-apa</td>
<td>‘anything’</td>
</tr>
</tbody>
</table>

The full reduplication of Malay nouns such as buŋa ‘flower’ usually indicates plurality, though there are exceptions like oɾɑŋ ‘person’, which, when reduplicated, results in a new lexical item oɾɑŋ-oɾɑŋ ‘scarecrow’. Reduplicating an adjective usually results in a change of category. For example, lambat ‘slow’ reduplicates to become the adverb lambat-lambat ‘slowly’. Reduplicating a wh-word such as apa ‘what’ results in the indefinite apa-apa ‘anything’.

In Malay, the voicing of the stem-initial stop usually determines the phonological outcome of məN-prefixation, which has many syntactic functions when it attaches to verb stems (ref. Soh & Nomoto 2011). Given that Malay does not permit NC (nasal/voiceless stop) clusters or non-homorganic NC (nasal/voiced stop) clusters, these illicit sequences formed through the prefixation process are repaired by nasal substitution and nasal assimilation respectively (e.g. Hassan 1974, Farid 1980). Nasal substitution refers to the process whereby the stem-initial voiceless stop coalesces with the nasal in the prefix to produce a nasal that has the same place of articulation as the stop (e.g. Pater 2001).3 On the other hand, nasal assimilation refers to the process in which the nasal in the prefix assimilates in place to the stem-initial stop but does not replace it. Examples of these two processes are given in (3).

(3) a. Nasal substitution when stem begins with a voiceless stop

/məN+pukul/  mənukul ‘hit’
/məN+tari/   mənari ‘dance’
/məN+kədʒər/ məŋədʒər ‘chase’

b. Nasal assimilation when stem begins with a voiced stop

/məN+bunoh/  məmbunoh ‘kill’
/məN+duɡa/   mənduɡa ‘suspect’
/məN+ɡanti/  məŋɡanti ‘change’

2 The examples given here are all disyllabic because there are few native Malay words that have more than two syllables. We maintain that Malay exhibits whole-word reduplication and not merely foot reduplication because longer loanwords, e.g. telefon ‘telephone (sg.)’, are also reduplicated as a whole, i.e. telefon-telefon ‘telephone (pl.)’.

3 In this paper, we assume Pater’s (2001) analysis of nasal substitution, in which the final nasal of the prefix and the initial voiceless stop of the stem coalesce into (i.e. are substituted by) a single nasal segment that is homorganic to the stop. However, we remain agnostic between this analysis and others, for example those in which the final nasal of the prefix assimilates to the place of the initial stop, then the initial stop deletes (e.g. Onn 1980).
Examples in (3a) show the phonological output of məN- prefixation when the stem begins with a voiceless stop. In these cases, the final nasal of the məN- prefix and the initial stop are both replaced by a nasal that is homorganic to the stop. For example, when məN- attaches to the stem pukul, we get a bilabial nasal, /m/, at the morpheme boundary. In contrast, examples in (3b) show the output of məN- prefixation when the stop at the beginning of the stem begins with a voiced stop. For these words, the final nasal in məN- takes the place feature of the following stem-initial stop, and the stop is still pronounced. For example, when məN- attaches to the stem bunʊh, we get the sequence /mb/ at the morpheme boundary.

2.2 Verbal reduplication with məN- prefix
Under verbal reduplication, the məN- prefix can attach to either the first or the second component of the reduplicated word. When məN- appears on the first component, the reduplicated word carries a ‘continuity’ meaning; when məN- appears on the second component, the reduplicated word carries a reciprocal meaning.

When the stem begins with a voiced stop, the copies in the reduplicated form with məN- prefixation are identical since the addition of the prefix does not change the form of the stem. Examples are given below in (4).

(4) Verbal reduplication when stem begins with voiced stop

<table>
<thead>
<tr>
<th>məN+X-X</th>
<th>X-məN+X</th>
<th>Stem</th>
<th>Stem gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>məmbunoh-bunoh</td>
<td>bunoh-məmbunoh</td>
<td>bunoh</td>
<td>‘kill’</td>
</tr>
<tr>
<td>məndu-ga-du-ga</td>
<td>du-ga-məndu-ga</td>
<td>du-ga</td>
<td>‘suspect’</td>
</tr>
<tr>
<td>məŋanti-ganti</td>
<td>ganti-məŋanti</td>
<td>ganti</td>
<td>‘change’</td>
</tr>
</tbody>
</table>

From the examples in (4), we can see that regardless of whether məN- attaches to the first or second component of the reduplicated verb, both components are pronounced the same. However, when the root word begins with a voiceless stop, the copies in the reduplicated form with məN- prefixation are no longer identical, as shown in (5).

(5) Verbal reduplication when stem begins with voiceless stop

<table>
<thead>
<tr>
<th>məN+X-X</th>
<th>X-məN+X</th>
<th>Stem</th>
<th>Stem gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>məmukul-mukul</td>
<td>pukul-məmukul</td>
<td>pukul</td>
<td>‘hit’</td>
</tr>
<tr>
<td>mənari-nari</td>
<td>tari-mənari</td>
<td>tari</td>
<td>‘dance’</td>
</tr>
<tr>
<td>məndʒar-ŋar</td>
<td>kədʒar-məndʒar</td>
<td>kədʒar</td>
<td>‘chase’</td>
</tr>
</tbody>
</table>

When məN- attaches to the first component, nasal substitution occurs on the first component. Additionally, the initial segment on the second component also surfaces as the homorganic nasal (e.g. mənari-nari instead of *mənari-tari), even though the context for nasal substitution is not met, thus constituting a case of overapplication. When məN- attaches to the second component, nasal substitution applies normally, only occurring on the second component, while the first component continues to have as its initial consonant a voiceless stop (e.g. tari-mənari instead of *nari-mənari). In the next section, we discuss the difficulties in accounting for this pattern of nasal substitution.

3 Other Accounts of Reduplication with məN- Prefixation

3.1 Derivational accounts
In any derivational account of phonology, the underlying representation along with a set of phonological processes contain all the information necessary to generate the appropriate surface form. Our case involves two morphological processes: full reduplication and məN- prefixation. The full reduplication of the verb

---

4 Other examples in which the form of the reduplicant varies as a function of its position relative to other morphology can be found in Lushootseed (Urbanczyk 1996, 2006), Kosraean (Kennedy 2005), and Mandarin (Lee-Kim 2016).
involves copying the phonological material of the root verb in its entirety. Affixing the məN- prefix to a stop-initial verb stem creates an illicit sequence at the morpheme boundary, which must be repaired by the phonological process of either nasal assimilation or nasal substitution.

In the case of verbs with roots beginning with a voiced stop, applying the two processes in either order yields the same surface form since the form of the stem does not change. Examples of both orders of process application on the verb bunʊh ‘kill’ are given in (6).

(6) a. Reduplication > məN- prefixation

<table>
<thead>
<tr>
<th>Process</th>
<th>‘continuous’</th>
<th>reciprocal</th>
</tr>
</thead>
<tbody>
<tr>
<td>məN+X-X</td>
<td>X-məN+X</td>
<td></td>
</tr>
<tr>
<td>reduplication</td>
<td>bunʊh-bunʊh</td>
<td>bunʊh-bunʊh</td>
</tr>
<tr>
<td>məN- prefixation</td>
<td>məmbunʊh-bunʊh</td>
<td>bunʊh-məmbunʊh</td>
</tr>
</tbody>
</table>

b. məN prefixation > Reduplication

<table>
<thead>
<tr>
<th>Process</th>
<th>‘continuous’</th>
<th>reciprocal</th>
</tr>
</thead>
<tbody>
<tr>
<td>məN+X-X</td>
<td>X-məN+X</td>
<td></td>
</tr>
<tr>
<td>məN- prefixation</td>
<td>məmbunʊh</td>
<td>məmbunʊh</td>
</tr>
<tr>
<td>reduplication</td>
<td>məmbunʊh-bunʊh</td>
<td>bunʊh-məmbunʊh</td>
</tr>
</tbody>
</table>

In (6a), reduplication is shown applying before məN- prefixation. The reduplication process produces a form with two copies of the verb. The məN- prefix is then free to attach to either copy with no phonological consequence to the stem. In (6b), the order of reduplication and məN- prefixation is reversed, with the same outcome. The məN- prefix first attaches to the verb stem, and the nasal assimilates to the place of the following stop, but the stop remains unchanged. The phonological material from the unchanged verb stem is then copied in the reduplication process. However, when the root verb begins with a voiceless stop, neither order of the processes will derive the correct result simultaneously for both the continuity meaning, where məN- affixes to the first component, and the reciprocal meaning, where məN- affixes to the second component.

(7) a. Reduplication > məN- prefixation

<table>
<thead>
<tr>
<th>Process</th>
<th>‘continuous’</th>
<th>reciprocal</th>
</tr>
</thead>
<tbody>
<tr>
<td>məN+X-X</td>
<td>X-məN+X</td>
<td></td>
</tr>
<tr>
<td>reduplication</td>
<td>tari-tari</td>
<td>tari-tari</td>
</tr>
<tr>
<td>məN- prefixation</td>
<td>*mənarι-tari</td>
<td>tari-mənarι</td>
</tr>
</tbody>
</table>

b. məN- prefixation > Reduplication

<table>
<thead>
<tr>
<th>Process</th>
<th>‘continuous’</th>
<th>reciprocal</th>
</tr>
</thead>
<tbody>
<tr>
<td>məN+X-X</td>
<td>X-məN+X</td>
<td></td>
</tr>
<tr>
<td>məN- prefixation</td>
<td>mənarι</td>
<td>mənarι</td>
</tr>
<tr>
<td>reduplication</td>
<td>mənarι-nari</td>
<td>*nari-mənarι</td>
</tr>
</tbody>
</table>

If reduplication occurs before məN- prefixation, as in (7a), then the voiceless stop would be copied in the reduplicant. Thus when məN- is affixed, only the component to which it is attached will undergo nasal substitution while the other non-prefixed component will always surface with the underlying voiceless stop. The result is that the correct form is derived for the word with reciprocal meaning in which məN- attaches to the second component, but the incorrect form is derived for the word with the continuous meaning in which məN- attaches to the first component. If məN- prefixation occurs before reduplication, as in (7b), the initial voiceless stop on the stem would be substituted by a homorganic nasal. When the stem is then reduplicated, both copies will begin with a nasal, thus deriving the correct form for the word in which məN- attaches to the first component, but not for the word in which məN- attaches to the second component. Clearly, a derivational
account like this cannot account for the forms of reduplicated verbs with the continuous and reciprocal meanings.

Given that \( məN- \) prefixation on the first or second component of a reduplicated verb results in semantic differences, we might also entertain the possibility that the order of the two morphological processes could be different for the continuous meaning and the reciprocal meaning. For the continuous meaning, prefixation could precede reduplication, as in (8).

(8) Morphosyntactic structure for \( mənari-nari \) ‘dance (continuous)’

As a first step in (8), \( məN- \) attaches to the root \( tari \) forming \( mənari \). Then, \( tari \) reduplicates and the reduplicant attaches to \( mənari \), yielding the correct surface form \( mənari-nari \). From a morphosyntactic perspective, this order is unproblematic if we assume a left-branching structure.

For the reciprocal meaning, reduplication must precede prefixation to yield the correct surface form \( tari-mənari \). As can be seen in (9), \( tari \) is reduplicated to form \( tari-tari \). Then, \( məN- \) must attach to the second component of the reduplicant.

(9) Morphosyntactic structure for \( tari-mənari \) ‘dance (reciprocal)’

However, this later morphological process of prefixation would require us to merge the prefix between two structures that have already been built. This is problematic in current models of morphosyntax. To avoid the insertion of new structure between the two copies of the reduplicated verb, we would have to stipulate that \( mən- \) merges to the left of \( tari-tari \), forming \( mənari-tari \), then the second copy of the reduplicant moves to the left of \( mənari \) to derive the correct surface order. Movement in syntax is typically motivated by the need to check certain features. In this case, it is difficult to argue that the movement of the reduplicant is motivated by anything other than the need to derive the correct surface order.

In this section we examined two different derivational analyses. In the first analysis, we assumed that the order of application for the two morphological processes, reduplication and \( məN- \) prefixation, was the same for whether the prefix attaches to the first or the second component of the reduplicated verb. This type of analysis was able to derive the correct surface form for one of the two types of verbs but not both. In the second analysis, we allowed for each type of verb to have a different order of application for the two processes. However, we would either have to violate basic syntactic assumptions, or stipulate unnecessary movement. Thus, neither derivational accounts adequately capture the phenomenon described. Therefore, following McCarthy and Prince (1995), we adopt a parallel Optimal Theoretic approach in which all the necessary morphology is completed first, and sound patterns are dealt with later in the phonology. We first give their analysis of the nasal substitution patterns for reduplicated verbs beginning with voiceless stops, its successes, and potential issues. Then, in Section 4, we present our revised OT analysis.

3.2 Parallel OT

In Optimality Theory (Prince & Smolensky 1993), the phonological output is determined by the interactions between faithfulness constraints comparing the input and output strings and markedness constraints on surface
configurations. McCarthy and Prince (1995) analyse precisely the məN- prefixation and reduplication data set. They account for the over/underapplication problems in reduplication by invoking the notion of base-reduplicant faithfulness to militate against differences between the two copies of the reduplicant. The three types of cover constraints used in their analysis are defined in (10).

(10) Constraints from McCarthy and Prince’s (1995) analysis

a. NASALSUB: A cover constraint for finer constraints that result in the substitution of the final nasal in məN- and the initial voiceless stop in the stem by a nasal homorganic to the stop.\(^5\) One violation is incurred for every NČ sequence in a candidate.
b. FAITH-IO: The cover constraint for all input-output faithfulness constraints. One violation is incurred for every mismatch between a segment in the input and a segment in the base of the output.
c. FAITH-BR: The cover constraint for all base-reduplicant faithfulness constraints. One violation is incurred for every mismatch between a segment in the base of a candidate and a reduplicant of the same candidate.

In the tableaux below, we present an adapted version of McCarthy and Prince’s OT analysis of the verbal reduplication phenomenon with these constraints. The base and reduplicant in the candidate set are marked by subscript B and R respectively.

(11) OT analysis, adapted from McCarthy and Prince (1995)
a. məN- attaches to first component

<table>
<thead>
<tr>
<th>/məN+tari-RED/</th>
<th>NASALSUB</th>
<th>FAITH-IO</th>
<th>FAITH-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. məntari_B-tari_R</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. mənari_B-nari_R</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. mənari_B-tari_R</td>
<td>*</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

b. məN- attaches to second component

<table>
<thead>
<tr>
<th>/tari-məN+RED/</th>
<th>NASALSUB</th>
<th>FAITH-IO</th>
<th>FAITH-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tari_B-məntari_R</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. nari_B-mənari_R</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. tari_B-mənari_R</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

In both tableaux in (11), we can see that NASALSUB has to outrank FAITH-IO to ensure that the candidate in which nasal substitution occurs wins over the faithful candidate. In (11a) when the prefix attaches to the first component, i.e. the base, candidates (b) and (c) incur the same number of violations of NASALSUB and FAITH-IO, but candidate (c) also incurs a violation of FAITH-BR, and is therefore harmonically bound by candidate (b). In (11b) when the prefix attaches to the second component, i.e. the reduplicant, candidate (c) no

\(^5\) We will not discuss in detail the constraints involved in nasal substitution since there are multiple possible analyses, and the specifics are not crucial to the current proposal. For one possible analysis of nasal substitution in məN-prefixation, we refer you to Pater (2001).
longer violates \textit{FAITH-IO} since the base is faithful to the input. Thus \textit{FAITH-IO} must outrank \textit{FAITH-BR}, allowing candidate (c) to win over candidate (b).

McCarthy and Prince’s analysis is successful in capturing both patterns of the verbal reduplication with a single ranking of constraints. However, their analysis crucially relies on specifying which component is the base and which is the reduplicant. In this case of reduplication, the first component must be the base and the second component the reduplicant. Flipping the base and reduplicant in the case of the reciprocal meaning results in the wrong output being chosen, as shown in the tableau in (12).

(12)

<table>
<thead>
<tr>
<th>/RED-məN+tari/</th>
<th>NASALSUB</th>
<th>FAITH-IO</th>
<th>FAITH-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tari_r-monari_B</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. nari_r-monari_B</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. tari_r-monari_B</td>
<td>*</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

We argue that this assumption about base and reduplicant is arbitrary and does not reflect speakers’ knowledge about base and reduplicant identity. We asked three native speakers of Malay for their judgments on base/reduplicant identity in various types of reduplication. In some cases, these native speakers were able to determine which part of the reduplicated word is the base and which is the reduplicant, while in other cases they were not. For example, in the case of rhythmic reduplication, e.g. \textit{kwe-nwe} ‘pastries’, our speakers identified the second component, \textit{nwe}, as the reduplicant as it does not stand as a word on its own. Similarly, in the case of \textit{bolak-balik} ‘back and forth’, \textit{bolak} is identified as the reduplicant because it does not exist as a stand-alone lexical item. Note here that order does not determine base- or reduplicanthood since there are words in which the reduplicant is the first component and words in which the reduplicant is the second component. Contrastively, in simple full reduplication, e.g. \textit{buŋa-buŋa} ‘flowers’, speakers were unable to identify one copy as the base and the other as the reduplicant since the two copies are identical.

In the case of the verbal reduplication with \textit{məN}-prefixation, speakers are able to distinguish the base and reduplicant when \textit{məN}- is attached to the first component, but not when it is attached to the second component. For \textit{manari-nari} ‘dance (cont.)’, \textit{nari} is not a valid word, hence is the reduplicant. For \textit{tari-monari} ‘dance (recip.)’, both \textit{tari} and \textit{monari} are valid words, speakers were hesitant to identify either component as the reduplicant. The uncertainty speakers had in determining the base vs. reduplicant in the latter type of verbal reduplication indicates Malay speakers do not necessarily need to know which copy is the base and which is the reduplicant in order to produce the correct phonological output. Yet it is precisely in this case, where \textit{məN}-attaches to the second component, that the identity of the base is crucial for McCarthy and Prince’s analysis presented above. Thus, in the next section, we present an alternative OT analysis that does not require the potentially stipulative designation of the base and reduplicant, thereby more accurately reflecting speakers’ intuition about these reduplicated forms.

4 Our Analysis
We argue that the verbal reduplication pattern seen here is best accounted for using a revised parallel OT analysis. This analysis does not require us to assume that the first component in the reduplicated verb is the base. Instead, we introduce a new markedness constraint that draws on a phonotactic trait of Malay which is also evident in other Austronesian languages.

4.1 Drawing upon Malay Phonotactics
Surveying the Austronesian languages of Southeast Asia, Blust (2013) finds that initial nasals are rare. For example, root-initial nasals are disfavored in Tagalog (Zuraw 2010). Malay displays a similar dispreference for word initial nasals. In Table 1, we present Malay nasal counts by place of articulation (bilabial, alveolar,
and velar) and percentages within each place type based on word-type frequencies\(^6\) using data from the An Crúbadáin Malay Corpus (Scannell 2007). Red cells indicate overrepresentation and blue cells indicate underrepresentation, \(\chi^2(4, \ N=48759) = 12687.074, \ p < 0.00001\). For full results of the chi-square analysis, refer to the Appendix.

**Table 1:** Counts of nasals by place of articulation and percentages by position in word. Data from the An Crúbadáin Malay Corpus

<table>
<thead>
<tr>
<th></th>
<th>Initial (%)</th>
<th>Medial (%)</th>
<th>Final (%)</th>
<th>Total Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>40%</td>
<td>52%</td>
<td>9%</td>
<td>15,890</td>
</tr>
<tr>
<td>n</td>
<td>4%</td>
<td>55%</td>
<td>40%</td>
<td>25,564</td>
</tr>
<tr>
<td>ŋ</td>
<td>2%</td>
<td>64%</td>
<td>34%</td>
<td>7,305</td>
</tr>
</tbody>
</table>

As shown in Table 1, word-initial nasals are largely dispreferred with the exception of the bilabial /m/.\(^7\) This dispreference for word-initial nasals is further highlighted in the graph in (13), which displays nasal counts by position.

(13) Nasal counts by position in words

![Nasal counts by position in words](image)

Clearly, the number of word-initial nasals is far fewer than nasals in other positions. We draw on this phonotactic tendency in our revised analysis of verbal reduplication with \(məN\)-prefixation.

Recall that the problematic case for McCarthy and Prince’s OT analysis was *tari-monari* in which \(məN\)-attaches to the second component of the reduplicated verb. For this type of verbal reduplication, it was necessary to assume that the first component is the base and the second component is the reduplicant. In our revised analysis, we get rid of this arbitrary assumption. Instead, we propose that the first component surfaces with an initial stop because word-initial nasals as phonotactically dispreferred, and this dis preference is stronger than the preference for the base and reduplicant to be identical. However, this phonotactic constraint should not affect words in which the initial segment is a nasal in the input (e.g. \(məN\)-initial words), only those in which the word-initial nasal is derived. To capture this generalization, we keep the **NASALSUB** cover constraint, which drives nasal substitution in the context of a nasal followed by a voiceless stop. In addition, we introduce a markedness constraint that militates against word-initial nasals. Finally, we replaced the cover

---

\(^6\) We chose to use word type frequency rather than token frequency because previous research (e.g. Hayes & Wilson 2008) finds that generalizations drawn from type frequencies better maps onto the phonotactics of a language.

\(^7\) Several instances of ‘morpheme-initial m-’ in Austronesian languages are prefixes that have fossilized (Blust 2013).
input-output and base-reduplicant faithfulness constraints with more specific constraints that militate against changes to nasals. These constraints are defined in (14).

(14) New and revised constraints
   a. *N[word-initial]: No nasals word-initially. Assign one violation for every word-initial nasal segment.
   b. PRESERVE(NASAL)-IO: Nasals in the input must be nasals in the output. Assign one violation for every nasal segment in the input that does not correspond to a nasal segment in the output.
   c. IDENT(NASAL)-BR: Corresponding segments in the base and reduplicant must have the same [nasal] value. Assign one violation for every segment in the base that does not have the same [nasal] value in the corresponding segment in the reduplicant.

The markedness constraint *N[word-initial] captures the phonotactic dispreference for nasals to occur in word-initial position. The modified input-output faithfulness constraint, PRESERVE(NASAL)-IO, protects nasals that were in the input from being denasalized in the output. The final base-reduplicant faithfulness constraint militates against differences in the nasality of corresponding segments between the two copies of the reduplicated verb. Note that though we continue to use the term ‘base-reduplicant,’ the analysis we present here does not require us to index identify the reduplicated components as the base or reduplicant. The ranking of these constraints is as in (15).

(15) NasalSub, Preserve(nasal)-IO  >> *N[word-initial] >> Ident(nasal)-BR

In the following tableaux, we show that these constraints with this ranking are able to account for the pattern of verbal reduplication whether the base is the first component or the second component. The example in (16) shows the analysis for cases in which məN- attaches to the first component.

(16) məN- attaches to the first component
   a. Base as first component, reduplicant as second component

<table>
<thead>
<tr>
<th>/məN+tari-RED/</th>
<th>NASALSUB</th>
<th>PRES(n)-IO</th>
<th>*N[word-initial]</th>
<th>Id(n)-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. məntari-tari</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. mənari-nari</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. mənari-tari</td>
<td></td>
<td>*</td>
<td>*</td>
<td>!</td>
</tr>
<tr>
<td>d. bənari-nari</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. Reduplicant as the first component and the base as the second component

<table>
<thead>
<tr>
<th>/məN+RED-tari/</th>
<th>NASALSUB</th>
<th>PRES(n)-IO</th>
<th>*N[word-initial]</th>
<th>Id(n)-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. məntari-tari</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. mənari-nari</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. mənari-tari</td>
<td></td>
<td>*</td>
<td>*</td>
<td>!</td>
</tr>
<tr>
<td>d. bənari-nari</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**NASALSUB** must dominate **IDENT(NASAL)-BR** to allow for nasal substitution where the context is met, thus ruling out candidate (a). **PRETEND(NASAL)-IO** must out-rank **N[word-initial]** so that input nasals are protected, thus ruling out candidate (d). Comparing candidates (b) and (c), both incur a violation of **N[word-initial]** because of the word-initial [m], however candidate (c) incurs an extra violation because of the mismatch in nasality between the initial segment in the first and second components of the reduplicated verb. That is, the first component begins with a nasal while the second component begins with a stop.

When **məN**- attaches to the second component of the reduplicant, the same grammar produces the correct winning candidate, as seen in (17).

(17) **məN**- attaches to the second component

<table>
<thead>
<tr>
<th>/tari-məN+RED/</th>
<th>NASALSUB</th>
<th>PRES(n)-IO</th>
<th>*N[word-initial]</th>
<th>ID(n)-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tari-məntari</td>
<td>*!</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. nari-mənari</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. tari-mənari</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/RED-məN+tari/</th>
<th>NASALSUB</th>
<th>PRES(n)-IO</th>
<th>*N[word-initial]</th>
<th>ID(n)-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tari-məntari</td>
<td>*!</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. nari-mənari</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. tari-mənari</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In both (17a) and (17b), we see the crucial ranking between **N[word-initial]** and **IDENT(NASAL)-BR**. Candidate (b) is ruled out by the higher ranked **N[word-initial]** since it has a word-initial nasal, leaving candidate (c) as the optimal candidate even though there is a mismatch between the base and reduplicant incurring a violation of **IDENT(NASAL)-BR**. Again, the analysis produces the correct output form, regardless of whether, in the input, the first component or the second component is the base. This can be taken to be a case of morphological richness of the base.

If these patterns are indeed driven by the markedness constraint **N[word-initial]**, then Malay verbal reduplication with **məN**- prefixation presents a case of the Emergence of the Unmarked (TETU). That is, initial nasals are typically allowed to surface (i.e. are unmarked) in Malay, such as when the word begins with a nasal underlingly. However, nasals derived from other processes are not allowed when occurring word-initially. TETU effects are commonly found in reduplicative morphology (e.g. McCarthy & Prince 1994), loanword phonology (e.g. Gouskova 2001), and second language phonology (e.g. Broselow et al. 1998).

We believe that this analysis is better suited to the patterns of reduplication presented here than McCarthy and Prince’s (1995) original analysis as it i) does away with ungrounded assumptions about base-reduplicant identity, and ii) is grounded in the phonotactics of the language.

A caveat to all phonological analyses of this phenomenon is, of course, whether these patterns are productive or lexicalized, since the latter would obviate the need for any phonological machinery at all. A score against complete lexicalization would be evidence from a text corpus of Malay that there is variation in the production of existing forms. We could also test for productivity by wug-testing Malay speakers to see if and how they generalize these patterns to novel verbs. Since such corpus and experimental studies are not within the scope of this paper, we will not discuss them further here.
5 Conclusion
To summarize, Malay verbal reduplication with \( məN- \) prefixation presents a unique case of overapplication as well as normal application. This pattern is shown to be difficult to account for in any derivational analysis. A previous parallel OT analysis relies on the assumption that the base is the first component. However, this view is not supported by speaker intuitions about these forms. We build upon the previous OT analysis by introducing a new markedness constraint, \( *N_{\text{word-initial}} \), which draws from phonotactic tendencies in Malay and related languages. This new grammar does not require learners to know which is the base and which is the reduplicant, therefore allowing for a richer base. Our analysis calls for the re-examination of other reduplication processes in which the relative position of the base and reduplicant is assumed.

References

**Appendix**

Chi-square test for nasal counts by position in word and place of articulation. Notation in each cell: count (expected total) [chi-square statistic]

<table>
<thead>
<tr>
<th>Results</th>
<th>Initial</th>
<th>Medial</th>
<th>Final</th>
<th>Row Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
<td>6310</td>
<td>8200</td>
<td>1380</td>
<td>15890</td>
</tr>
<tr>
<td></td>
<td>(2448.45)</td>
<td>(8814.98)</td>
<td>(42.90)</td>
<td></td>
</tr>
<tr>
<td>Alveolar</td>
<td>1077</td>
<td>14146</td>
<td>10241</td>
<td>25554</td>
</tr>
<tr>
<td></td>
<td>(3935.87)</td>
<td>(14181.50)</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>Velar</td>
<td>120</td>
<td>4703</td>
<td>2482</td>
<td>7305</td>
</tr>
<tr>
<td></td>
<td>(1124.60)</td>
<td>(4052.44)</td>
<td>(0.44)</td>
<td></td>
</tr>
<tr>
<td><strong>Column Totals</strong></td>
<td>7507</td>
<td>27049</td>
<td>14203</td>
<td>48759 (Grand Total)</td>
</tr>
</tbody>
</table>

$\chi^2 (4, N=48759) = 12687.074, p < 0.00001$

Calculated using the Chi-Square Test Calculator (Stangroom 2018).